Chapter 2
Fredholm Integral Equations of the Second Kind (General Kernel)

In Chap. 1, we conducted a thorough examination of the Fredholm integral equation of the second kind for an arbitrary complex parameter λ, assuming that the free term $f(x)$ is complex-valued and continuous on the interval $[a,b]$ and that the kernel $K(x,t)$ is complex-valued, continuous, and separable on the square $Q(a,b) = \{(x,t)\colon [a,b] \times [a,b]\}$. We stated the four Fredholm theorems and the Fredholm Alternative Theorem which provide for the construction of all possible solutions to the equation under these assumptions.

A question naturally arises: What, if anything, can be proven if $K(x,t)$ is a *general* kernel, i.e., an arbitrary kernel that is only assumed to be complex-valued and continuous? In this chapter, we will answer this question completely by proving that all of the Fredholm theorems continue to hold in this eventuality.

In Sect. 2.1, we present several tools of the trade which are indispensible for the comprehension of the material in this chapter.

In Sects. 2.2 and 2.3, we use these tools to show that the Fredholm integral equation of the second kind with a general kernel has a unique solution if the product of the parameter λ and the "size" of the kernel is small.

In Sect. 2.4, we prove the validity of the Fredholm theorems for unrestricted λ and a general kernel.

In Sect. 2.5, we show how to construct the resolvent kernel that appears in the solution to the integral equation recursively.

In Sect. 2.6, we introduce numerical methods for producing an approximation to the solution of a Fredholm integral equation. These methods are necessary due to the inherent computational difficulties in constructing the resolvent kernel.

2.1 Tools of the Trade

In this chapter, the reader should be familiar with the following topics:

- *Norms of continuous functions*: Let $\mathcal{C}[a,b]$ denote the vector space that consists of all complex-valued continuous functions on the interval $[a,b]$. A *norm* on $\mathcal{C}[a,b]$

S.M. Zemyan, *The Classical Theory of Integral Equations: A Concise Treatment*,
DOI 10.1007/978-0-8176-8349-8_2, © Springer Science+Business Media, LLC 2012

is a real-valued function ($f \in \mathcal{C}[a,b] \to \|f\| \in \mathbb{R}$) with the following properties:

1. $\|f\| \geq 0$ for every $f(x) \in \mathcal{C}[a,b]$
2. $\|f\| = 0$ if and only if $f(x) \equiv 0$ on $[a,b]$
3. $\|kf\| = |k|\,\|f\|$
4. $\|f+g\| \leq \|f\| + \|g\|$

In particular, we will have a strong interest in the *p-norm* $(0 < p < +\infty)$ defined by

$$\|f\|_p = \left(\int_a^b |f(x)|^p \, \mathrm{d}x \right)^{1/p}$$

in the special cases $p = 1$ and $p = 2$. Of critical importance is the *Cauchy–Schwarz inequality* which states that

$$|\langle f, g \rangle| \leq \|f\|_2 \, \|g\|_2$$

or, more explicitly, that

$$\left| \int_a^b f(x)\overline{g(x)} \, \mathrm{d}x \right| \leq \left(\int_a^b |f(x)|^2 \, \mathrm{d}x \right)^{1/2} \left(\int_a^b |g(x)|^2 \, \mathrm{d}x \right)^{1/2}.$$

Another norm of interest is the *supremum norm* defined by

$$\|f\|_\infty = \sup\{\, |f(x)| \colon x \in [a,b] \,\}.$$

In an entirely similar way, a norm can be defined on the set of complex-valued continuous functions defined on the square $Q(a,b)$. Kernels, in particular, can be normed. The norms

$$\|K\|_2 = \left(\int_a^b \int_a^b |K(x,t)|^2 \, \mathrm{d}x \, \mathrm{d}t \right)^{1/2}$$

and

$$\|K\|_\infty = \sup\{\, |K(x,t)| \colon (x,t) \in Q(a,b) \,\}$$

will be of special interest in this chapter.

- *Uniform convergence of an infinite sequence of functions*: An infinite sequence $\{f_n(x)\}$ of functions *converges uniformly* on the interval $[a,b]$ to a function $f(x)$ if, for every $\varepsilon > 0$, there exists an integer $N = N(\varepsilon)$ such that $|f_n(x) - f(x)| < \varepsilon$ for all $x \in [a,b]$ and all $n \geq N(\varepsilon)$.

  An infinite series $\Sigma_1^\infty f_n(x)$ converges uniformly on $[a,b]$ if its sequence of partial sums converges uniformly on $[a,b]$.

  The *Cauchy criterion* is used to establish uniform convergence. We say that an infinite sequence $\{f_n(x)\}$ of functions defined on $[a,b]$ converges uniformly there if and only if, for every $\varepsilon > 0$, there exists a fixed integer $N(\varepsilon)$ such that $|f_n(x) - f_m(x)| < \varepsilon$ for all $x \in [a,b]$ and all $n, m \geq N(\varepsilon)$.

Uniform convergence is an essential hypothesis in many theorems. For example, if $\{f_n(x)\}$ is an infinite sequence of continuous functions on $[a,b]$ and if the sequence $\{f_n(x)\}$ converges uniformly to the limit function $f(x)$ there, then $f(x)$ is continuous on $[a,b]$.

Uniform convergence is also required to justify term-by-term integration. If $\{f_n(x)\}$ is a sequence of integrable functions that converge uniformly to $f(x)$ on $[a,b]$, then $f(x)$ is integrable and

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x.$$

As an immediate consequence, we can say that if

$$f(x) = \sum_{n=1}^{\infty} f_n(x)$$

and the convergence is uniform on the interval $[a,b]$, then

$$\int_a^b f(x)\,\mathrm{d}x = \sum_{n=1}^{\infty} \int_a^b f_n(x)\,\mathrm{d}x.$$

- *Analytic, entire, and meromorphic functions*: A *region* is a nonempty, connected, open subset of the plane. If $f(z)$ is a complex-valued function defined in the region Ω and $f(z)$ is differentiable for every $z_0 \in \Omega$, then we say that $f(z)$ is *analytic* or *holomorphic* in Ω.

  Suppose that $f_1(z)$ is analytic on Ω_1 and that $f_2(z)$ is analytic on Ω_2. If $\Omega_1 \cap \Omega_2 \neq \emptyset$ and $f_1(z) = f_2(z)$ for all $z \in \Omega_1 \cap \Omega_2$, then $f_1(z)$ and $f_2(z)$ are said to be *direct analytic continuations* of one another.

  A function that is analytic in the whole complex plane $\mathbb{C}$ is called *entire*.

  A function $f(z)$ is *meromorphic* in a region Ω if there is a set P such that (a) P has no limit point in Ω, (b) $f(z)$ is analytic in the region $\Omega \setminus P$, and (c) $f(z)$ has a pole at each point of P. Every meromorphic function in a region Ω is a quotient of two functions which are analytic in Ω.

2.2 The Method of Successive Substitution

Consider the integral equation

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t,$$

where we assume as usual that $f(x)$ is continuous on the interval $[a,b]$ and that $K(x,t)$ is complex-valued and continuous on the square $Q(a,b)$. If this equation has a solution $\phi(x)$, then the equation itself provides a representation for it. Without any

preconditions other than its integrability, $\phi(x)$ can be substituted into the integrand, thereby producing yet another, albeit more complicated, representation for $\phi(x)$ in terms of itself. More precisely, if we replace x by t and t by s in the integral equation, then a direct substitution yields

$$\begin{aligned}\phi(x) &= f(x) + \lambda \int_{t=a}^{b} K(x,t) \left[f(t) + \lambda \int_{s=a}^{b} K(t,s)\,\phi(s)\,\mathrm{d}s \right] \mathrm{d}t \\ &= f(x) + \lambda \int_a^b K(x,t)\,f(t)\,\mathrm{d}t + \lambda^2 \int_{t=a}^{b} \int_{s=a}^{b} K(x,t)\,K(t,s)\,\phi(s)\,\mathrm{d}s\,\mathrm{d}t.\end{aligned}$$

After interchanging the order of integration in the last integral and replacing the dummy variable s with t, we have

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,f(t)\,\mathrm{d}t + \lambda^2 \int_a^b K_2(x,t)\,\phi(t)\,\mathrm{d}t,$$

where we have set

$$K_2(x,t) = \int_a^b K(x,s)\,K(s,t)\,\mathrm{d}s.$$

It may seem pointless to repeat this process, since the solution $\phi(x)$ will always be represented in terms of itself, no matter how many times we repeat it. In actuality, however, successive substitution proves to be quite fruitful. Not only does the continuation of this iterative process eventually produce the solution to the equation, but it also produces an elegant representation for it that does not involve $\phi(x)$. Additional iterations lead to the general form

$$\phi(x) = f(x) + \sum_{m=1}^{n} \lambda^m \left(\int_a^b K_m(x,t)\,f(t)\,\mathrm{d}t \right) + \lambda^{n+1} \int_a^b K_{n+1}(x,t)\,\phi(t)\,\mathrm{d}t,$$

for any integer n, where initially we have set $K_1(x,t) = K(x,t)$, and then

$$K_m(x,t) = \int_a^b K_{m-1}(x,s)\,K(s,t)\,\mathrm{d}s, \tag{2.1}$$

for each $m = 2,\ldots,n$. This general form is valid for all λ.

The functions $K_m(x,t)$ are called *iterated kernels*. Each $K_m(x,t)$ is complex-valued and continuous on $Q(a,b)$. Also, $K_m(x,t)$ is bounded there for all $m \geq 2$, for if $|K(x,t)| \leq M$, then $|K_m(x,t)| \leq M^m\,(b-a)^{m-1}$.

Now let

$$\sigma_n(x) = \sum_{m=1}^{n} \lambda^{m-1} \left(\int_a^b K_m(x,t)\,f(t)\,\mathrm{d}t \right) \tag{2.2}$$

and

$$\rho_n(x) = \lambda^{n+1} \int_a^b K_{n+1}(x,t)\,\phi(t)\,\mathrm{d}t,$$

so that

$$\phi(x) = f(x) + \lambda\, \sigma_n(x) + \rho_n(x).$$

With a little effort, we show that the sequence $\{\sigma_n(x)\}$ of continuous functions converges uniformly to a continuous limit function $\sigma(x)$ on the interval $[a,b]$ and that $\phi(x) = f(x) + \lambda\, \sigma(x)$.

Given the bound on $K_m(x,t)$, each term of the sum $\sigma_n(x)$ satisfies the inequality

$$\left| \lambda^{m-1} \left(\int_a^b K_m(x,t)\, f(t)\, \mathrm{d}t \right) \right| \leq (\, |\lambda|\, M\, (b-a)\,)^{m-1}\, M\, \|f\|_1.$$

If $|\lambda|\, M\, (b-a) < 1$, then the sequence $\{\sigma_n(x)\}$ of partial sums is a Cauchy sequence. For if $\varepsilon > 0$ is arbitrary, then

$$\begin{aligned} |\sigma_n(x) - \sigma_p(x)| &\leq \left[\sum_{m=p+1}^{n} (\, |\lambda|\, M\, (b-a)\,)^{m-1} \right] M\, \|f\|_1 \\ &\leq (|\lambda|\, M\, (b-a))^p\, \frac{M \|f\|_1}{1 - |\lambda|\, M\, (b-a)} \\ &< \varepsilon \end{aligned}$$

if p is large enough. In addition, the remainder term $\rho_n(x) \to 0$ uniformly on $[a,b]$ as $n \to +\infty$, in view of the estimate $|\rho_n(x)| \leq |\lambda|\, M\, \|\phi\|_1\, (|\lambda|\, M\, (b-a))^n$.

It now follows that the sequence $\{\sigma_n(x)\}$ of continuous functions converges absolutely and uniformly on the interval $[a,b]$ to the continuous limit function

$$\sigma(x) = \sum_{m=1}^{\infty} \lambda^{m-1} \left(\int_a^b K_m(x,t)\, f(t)\, \mathrm{d}t \right),$$

provided that $|\lambda|\, M\, (b-a) < 1$. Furthermore, since term-by-term integration is permitted as a consequence of uniform convergence, we have

$$\sigma(x) = \int_a^b \left(\sum_{m=1}^{\infty} \lambda^{m-1}\, K_m(x,t) \right) f(t)\, \mathrm{d}t = \int_a^b R(x,t;\lambda) f(t)\, \mathrm{d}t,$$

where $R(x,t;\lambda)$ denotes the infinite series in parentheses. This series is known as the *Neumann series*, and it is the *resolvent kernel* of the integral equation. Its radius of convergence is at least $1/(M(b-a))$.

Recall that a function $\phi(x)$ is a *solution* to an integral equation if the integral equation is transformed into an identity by the substitution of $\phi(x)$ into it.

We shall show that $f(x) + \lambda\, \sigma(x)$ is the solution to the integral equation by showing that

$$f(x) + \lambda\, \sigma(x) = f(x) + \lambda \int_a^b K(x,t)\, (f(t) + \lambda\, \sigma(t))\, \mathrm{d}t.$$

By substituting the series expansion for $\sigma(x)$ into the integral on the right, we obtain

$$\begin{aligned}
&\lambda \int_a^b K(x,t)\,[f(t)+\lambda\,\sigma(t)]\,\mathrm{d}t \\
&\quad = \lambda \int_a^b K(x,t)\,f(t)\,\mathrm{d}t \\
&\qquad + \lambda^2 \int_{t=a}^b K(x,t) \left[\int_{s=a}^b \left(\sum_{m=1}^{\infty} \lambda^{m-1} K_m(t,s) \right) f(s)\,\mathrm{d}s \right] \mathrm{d}t \\
&\quad = \lambda \int_a^b K(x,t)\,f(t)\,\mathrm{d}t \\
&\qquad + \lambda^2 \int_{s=a}^b \sum_{m=1}^{\infty} \lambda^{m-1} \left(\int_{t=a}^b K_m(x,t)\,K(t,s)\,\mathrm{d}t \right) f(s)\,\mathrm{d}s \\
&\quad = \lambda \int_a^b K(x,t)\,f(t)\,\mathrm{d}t + \lambda \int_a^b \left(\sum_{m=1}^{\infty} \lambda^{m} K_{m+1}(x,s) \right) f(s)\,\mathrm{d}s \\
&\quad = \lambda\,\sigma(x).
\end{aligned}$$

We can now rightfully claim that $\phi(x) = f(x) + \lambda\,\sigma(x)$. Since $f(x)$ is assumed to be continuous on the interval $[a,b]$ and we have proven that $\sigma(x)$ is continuous there, it follows that $\phi(x)$ is continuous as well and that its norm $\|\phi\|_1$ is finite.

The discussion above serves to establish the following result:

Theorem 2.2.1 (Successive Substitution). *Let $f(x)$ be a complex-valued and continuous function defined on the interval $[a,b]$, and let $K(x,t)$ be a complex-valued continuous kernel defined on the square $Q(a,b)$ and also bounded there by M. Let λ be a complex parameter. If $|\lambda|\,M\,(b-a) < 1$, then the unique solution to the Fredholm integral equation*

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t$$

is given by

$$\phi(x) = f(x) + \lambda \int_a^b R(x,t;\lambda)\,f(t)\,\mathrm{d}t,$$

where $R(x,t;\lambda)$ is the resolvent kernel

$$R(x,t;\lambda) = \sum_{m=1}^{\infty} \lambda^{m-1}\,K_m(x,t).$$

Section 2.2 Exercises

1. For each of the kernels $K(x,t)$ given below, determine all of the iterated kernels $K_m(x,t)$ and then form the resolvent kernel $R(x,t;\lambda)$. Note that the resolvent depends upon the endpoints of the interval $[a,b]$. For which values of λ does the resolvent series converge? Can you sum the resolvent series in closed form?

   (a) $A(x,t) = x+t$, defined on $Q(-1,+1)$
   (b) $B(x,t) = x-t$, defined on $Q(0,1)$
   (c) $C(x,t) = xt$, defined on $Q(a,b)$
   (d) $D(x,t) = (1+x)(1-t)$, defined on $Q(-1,0)$
   (e) $E(x,t) = x^2t^2$, defined on $Q(-1,+1)$
   (f) $F(x,t) = \cos x \cos t$, defined on $Q(0,\pi)$
   (g) $G(x,t) = \sin(x+t)$, defined on $Q(0,\pi)$
   (h) $H(x,t) = x\mathrm{e}^t$, defined on $Q(0,1)$
   (i) $I(x,t) = \mathrm{e}^{x-t}$, defined on $Q(0,1)$
   (j) $J(x,t) = \mathrm{e}^x \cos t$, defined on $Q(0,\pi)$

2. Suppose that the kernel $K(x,t)$ is continuous on the square $Q(a,b)$. Prove that if the kernel $K(x,t)$ is separable, then every iterated kernel $K_m(x,t)$, $m = 2,3,\ldots,$ is also separable.
3. Suppose that the kernel $K(x,t)$ is a function of the product xt, say $K(x,t) = G(xt)$. Is the iterated kernel $K_2(x,t)$ also a function of the product xt?
4. Suppose that the kernel $K(x,t)$ is a function of the difference $x-t$, say $K(x,t) = H(x-t)$. Is the iterated kernel $K_2(x,t)$ also a function of the difference $x-t$?
5. Is every kernel an iterated kernel? In other words, if the kernel $K(x,t)$ is continuous on the square $Q(a,b)$, does there exist a continuous kernel $L(x,t)$ such that
$$K(x,t) = \int_a^b L(x,s)\,L(s,t)\,\mathrm{d}s?$$
6. Let $m \geq 2$. Show that
$$K_m(x,t) = \int_a^b K_r(x,s)\,K_{m-r}(s,t)\,\mathrm{d}s$$
for every $r = 1,\ldots,m-1$.
Hint: If $m = 2$, then the conclusion holds by definition. Use an inductive argument for $m \geq 3$.

2.3 The Method of Successive Approximation

In this section, we introduce another recursive method for solving Fredholm integral equations of the second kind. There are some distinct advantages in considering another method: additional insight is gained into the recursive process, a different

proof of convergence is used, and a better result is obtained, in the sense that the resolvent series may have a larger radius of convergence.

Consider the Fredholm integral equation

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t. \tag{2.3}$$

If $\lambda = 0$, then $\phi(x) = f(x)$ is the unique solution to the equation. In view of this observation, it seems reasonable to suggest that if $|\lambda| \approx 0$, then $\phi(x) \approx f(x)$, i.e., that the free term $f(x)$ serves as a reasonable zeroth-order approximation $\phi_0(x)$ to the solution $\phi(x)$. However, for computational reasons, $f(x)$ may not be the most practical choice for larger values of λ. Depending upon the form of the kernel, the choices $\phi_0(x) = 1$ or $\phi_0(x) = \mathrm{e}^x$ or some other choice might be more reasonable from the perspective of integrability.

After $\phi_0(x)$ is chosen, a first-order approximation $\phi_1(x)$ to $\phi(x)$ is furnished by substituting $\phi_0(t)$ for $\phi(t)$ into the integrand to obtain

$$\phi_1(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi_0(t)\,\mathrm{d}t.$$

If $\phi_1(x) = \phi_0(x)$, then $\phi(x) = \phi_1(x)$, and the solution has been found. If the integral vanishes, then $\phi_1(x) = f(x)$. At this point, the iterative process continues as if we had originally chosen $\phi_0(x) = f(x)$.

If $\phi_1(x) \neq \phi_0(x)$, then the substitution of $\phi_1(t)$ into the integrand yields the second-order approximation

$$\phi_2(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi_1(t)\,\mathrm{d}t.$$

If $\phi_2(x) = \phi_1(x)$, then $\phi(x) = \phi_2(x)$, and the solution has been found. If the integral vanishes, then $\phi_2(x) = f(x)$. Again, the iterative process continues as if we had originally chosen $\phi_0(x) = f(x)$. If $\phi_2(x) = \phi_0(x)$, then a continuation of this iterative process produces the two distinct constant subsequences, namely, $\phi_0(x) = \phi_2(x) = \phi_4(x) = \cdots$ and $\phi_1(x) = \phi_3(x) = \phi_5(x) = \cdots$. If this were to happen, then a unique solution to the integral equation obviously would not exist.

If $\phi_2(x) \neq \phi_1(x)$ or $\phi_0(x)$, then a substitution of $\phi_2(t)$ into the integrand yields the third-order approximation

$$\phi_3(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi_2(t)\,\mathrm{d}t,$$

and comments similar to those made above regarding the two previous iterations can again be made.

Assuming that $\phi_n(x) \neq \phi_j(x)$ for $j = 0, \ldots, n-1$ and that $\phi(x) \neq f(x)$, then a substitution of $\phi_n(t)$ into the integrand yields the $(n+1)$st-order approximation

$$\phi_{n+1}(x) = f(x) + \lambda \int_a^b K(x,t)\, \phi_n(t)\, \mathrm{d}t.$$

Each approximant $\phi_n(x)$ has an alternate form. In Sect. 2.2, we substituted the integral equation into itself repeatedly; here, we substitute each approximant $\phi_j(x)$ into the expression for the next approximant $\phi_{j+1}(x)$ to obtain

$$\begin{aligned}\phi_{n+1}(x) = f(x) + \sum_{m=1}^{n} \lambda^m \left(\int_a^b K_m(x,t)\, f(t)\, \mathrm{d}t \right) \\ + \lambda^{n+1} \int_a^b K_{n+1}(x,t)\, \phi_0(t)\, \mathrm{d}t. \end{aligned} \tag{2.4}$$

Alternately, we have

$$\phi_{n+1}(x) = f(x) + \lambda\, \sigma_n(x) + \omega_{n+1}(x),$$

where $\sigma_n(x)$ was previously defined by (2.2), and we set

$$\omega_{n+1}(x) = \lambda^{n+1} \int_a^b K_{n+1}(x,t)\, \phi_0(t)\, \mathrm{d}t.$$

In Sect. 2.2, we showed that the sequence $\sigma_n(x)$ converges uniformly to the function $\sigma(x)$ on the interval $[a,b]$, provided that $|\lambda|$ was small enough relative to the size of the kernel. Here, we reprove this result under a different condition on λ.

An application of the Cauchy–Schwarz inequality to the definition of the iterated kernel yields

$$|K_m(x,t)|^2 \leq \left(\int_a^b |K_{m-1}(x,s)|^2\, \mathrm{d}s \right) \left(\int_a^b |K(s,t)|^2\, \mathrm{d}s \right).$$

Integrating this inequality with respect to t yields

$$\int_a^b |K_m(x,t)|^2\, \mathrm{d}t \leq \left(\int_a^b |K_{m-1}(x,s)|^2\, \mathrm{d}s \right) \left(\int_a^b \int_a^b |K(s,t)|^2\, \mathrm{d}s\, \mathrm{d}t \right),$$

for each fixed $x \in [a,b]$, or more simply

$$\kappa_m(x) \leq \kappa_{m-1}(x)\, \|K\|_2^2,$$

where we have set

$$\kappa_m(x) = \int_a^b |K_m(x,t)|^2\, \mathrm{d}t$$

and

$$\|K\|_2 = \left(\int_a^b \int_a^b |K(x,t)|^2 \, \mathrm{d}x \, \mathrm{d}t \right)^{1/2}.$$

Recursively, we easily obtain the estimate

$$\kappa_m(x) \le \kappa_1(x) \, \|K\|_2^{2m-2}.$$

Another application of the Cauchy–Schwarz inequality to the integrals in the sum $\sigma_n(x)$ provides the estimate

$$\begin{aligned} \left| \int_a^b K_m(x,t) \, f(t) \, \mathrm{d}t \right|^2 &\le \left(\int_a^b |K_m(x,t)|^2 \, \mathrm{d}t \right) \left(\int_a^b |f(t)|^2 \, \mathrm{d}t \right) \\ &= \kappa_m(x) \, \|f\|_2^2 \\ &\le \kappa_1(x) \, \|f\|_2^2 \, \|K\|_2^{2m-2}. \end{aligned}$$

Hence, each term in the sum $\sigma_n(x)$ can be estimated by the inequality

$$\left| \lambda^m \int_a^b K_m(x,t) \, f(t) \, \mathrm{d}t \right| \le \frac{\sqrt{\kappa_1(x)} \, \|f\|_2}{\|K\|_2} \, (|\lambda| \, \|K\|_2)^m, \tag{2.5}$$

which is valid for each fixed $x \in [a,b]$.

It follows from this estimate that the sequence $\sigma_n(x)$ converges absolutely and uniformly to a unique limit $\sigma(x)$ on the interval $[a,b]$ whenever $|\lambda| \, \|K\|_2 < 1$, since it is dominated by a convergent geometric series of positive terms.

By using a similar estimate, we also see that

$$|\omega_{n+1}(x)| \le \frac{\sqrt{\kappa_1(x)} \, \|\phi_0\|_2}{\|K\|_2} \, (|\lambda| \, \|K\|_2)^{n+1} \to 0$$

as $n \to +\infty$.

The results of the last two paragraphs imply that $\phi(x) = f(x) + \lambda \, \sigma(x)$.

The proof that $\phi(x)$ is unique proceeds by contradiction. Suppose that there were two distinct solutions to Eq. (2.3), say $\phi(x)$ and $\tilde{\phi}(x)$. If we set $\delta(x) = \phi(x) - \tilde{\phi}(x)$, then $\delta(x)$ satisfies the homogeneous integral equation

$$\delta(x) = \lambda \int_a^b K(x,t) \, \delta(t) \, \mathrm{d}t.$$

If $\delta(x) \neq 0$, then it would be an eigenfunction of the kernel corresponding to the eigenvalue λ. Showing that $\delta(x) \equiv 0$ not only proves that $\phi(x)$ is unique, but it also shows that the kernel has no eigenvalues smaller than $1/\|K\|_2$. An application of the Cauchy–Schwarz inequality gives

$$|\delta(x)|^2 \leq |\lambda|^2 \left(\int_a^b |K(x,t)|^2 \, \mathrm{d}t \right) \left(\int_a^b |\delta(t)|^2 \, \mathrm{d}t \right),$$

which can be integrated with respect to x yielding

$$\left(1 - |\lambda|^2 \|K\|_2^2\right) \int_a^b |\delta(x)|^2 \, \mathrm{d}x \leq 0.$$

The restriction $|\lambda| \, \|K\|_2 < 1$ implies that $\int_a^b |\delta(x)|^2 \, \mathrm{d}x = 0$.

The discussion above serves to establish the following result:

Theorem 2.3.1 (Successive Approximation). *Let λ be a complex parameter and let $f(x)$ be a complex-valued continuous function defined on the interval $[a,b]$. Let $K(x,t)$ be a complex-valued continuous kernel defined on the square $Q(a,b)$ with*

$$\|K\|_2 = \left(\int_a^b \int_a^b |K(x,t)|^2 \, \mathrm{d}x \, \mathrm{d}t \right)^{1/2}.$$

If $|\lambda| \, \|K\|_2 < 1$, then the unique solution to the Fredholm integral equation

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t) \, \phi(t) \, \mathrm{d}t$$

is given by

$$\phi(x) = f(x) + \lambda \int_a^b R(x,t;\lambda) \, f(t) \, \mathrm{d}t,$$

where $R(x,t;\lambda)$ is the resolvent kernel

$$R(x,t;\lambda) = \sum_{m=1}^{\infty} \lambda^{m-1} \, K_m(x,t).$$

The estimate for the radius of convergence in the Theorem of Successive Approximation is larger than the corresponding estimate in the Theorem of Successive Substitution. Indeed, $\|K\|_2 \leq M(b-a)$, with equality holding only if $K(x,t) \equiv M$. Hence,

$$\frac{1}{M(b-a)} \leq \frac{1}{\|K\|_2}.$$

For example, consider the kernel $K(x,t) = x^{10} t^{10}$ on the square $Q(0,1)$. For this kernel, we have $\|K\|_2 = \frac{1}{21}$ and $M(b-a) = 1$.

The inequalities established within the proof of the theorem may be used again to estimate the magnitude of the error incurred by using the approximant $\phi_n(x)$ to estimate $\phi(x)$. For each $x \in [a,b]$, we have

$$
\begin{aligned}
|\phi(x)-\phi_n(x)| &= \left| \sum_{m=n+1}^{\infty} \lambda^m \int_a^b K_m(x,t)\, f(t)\, \mathrm{d}t \right| \\
&\le \frac{\sqrt{\kappa_1(x)}\, \|f\|_2}{\|K\|_2} \sum_{m=n+1}^{\infty} (|\lambda|\, \|K\|_2)^m \\
&\le \frac{\|\sqrt{\kappa_1}\|_\infty \|f\|_2}{\|K\|_2 (1-|\lambda|\, \|K\|_2)} (|\lambda|\, \|K\|_2)^{n+1}.
\end{aligned}
$$

This inequality shows that if n is large enough, then the difference $|\phi(x)-\phi_n(x)|$ is uniformly small on the interval $[a,b]$. However, the rate of convergence may be quite slow if $|\lambda|\, \|K\|_2$ is close to 1. The implication here is that a large number of approximants may need to be computed, thereby decreasing the efficiency and convenience of the method.

Illustrative Examples

- *Example 1*: If it should happen that an iterated kernel $K_m(x,t) \equiv 0$ on the square $Q(a,b)$, then the resolvent kernel reduces to a polynomial in λ, and the kernel $K(x,t)$ will have no eigenvalues.

  In particular, if $K_2(x,t) \equiv 0$, then $R(x,t;\lambda) = K(x,t)$, and the kernel $K(x,t)$ is said to be *orthogonal to itself*. It is easy to invent such kernels in terms of trigonometric functions. For example, if $K(x,t) = \sin(x+2t)$, then $K_2(x,t) = 0$. Hence, the solution to the integral equation

$$\phi(x) = x + \int_0^{2\pi} \sin(x+2t)\, \phi(t)\, \mathrm{d}t$$

  is

$$\phi(x) = x + \int_0^{2\pi} \sin(x+2t)\, t\, \mathrm{d}t = x - \pi \cos x.$$

- *Example 2*: Consider the separable kernel

$$K(x,t) = \cos(x+t) = \cos x \cos t - \sin x \sin t$$

  defined on the square $Q(0,\pi)$. By employing the methods in Chap. 1, we can obtain the resolvent kernel in the form

$$R(x,t;\lambda) = -\frac{D(x,t;\lambda)}{D(\lambda)} = \frac{\cos x \cos t}{1-\frac{\lambda\pi}{2}} - \frac{\sin x \sin t}{1+\frac{\lambda\pi}{2}},$$

  showing that it is a meromorphic function of λ with simple poles at $\pm 2/\pi$.

Alternatively, the resolvent kernel can also be easily computed in terms of the iterated kernels of $K(x,t)$. Simple integrations yield

$$K_2(x,t) = \int_0^\pi \cos(x+s)\cos(s+t)\,\mathrm{d}s = \tfrac{\pi}{2}\cos(x-t)$$

and

$$K_3(x,t) = \int_0^\pi \cos(x-s)\cos(s+t)\,\mathrm{d}s = \left(\tfrac{\pi}{2}\right)^2 \cos(x+t) = \left(\tfrac{\pi}{2}\right)^2 K(x,t).$$

A short inductive argument shows that

$$K_m(x,t) = \begin{cases} \left(\frac{\pi}{2}\right)^{m-1}\cos(x+t), & \text{if } m \text{ is odd,} \\ \left(\frac{\pi}{2}\right)^{m-1}\cos(x-t), & \text{if } m \text{ is even.} \end{cases}$$

After inserting the iterated kernels into the resolvent series and simplifying the result, we obtain the same expression for $R(x,t;\lambda)$ as above.

The Theorem of Successive Substitution guarantees that the resolvent series converges if $|\lambda| \le 1/M(b-1) < 1/\pi$; the Theorem of Successive Approximation guarantees convergence if $|\lambda| \le 1/\|K\|_2 < \sqrt{2}/\pi$. If we were to expand $R(x,t;\lambda)$ in a Taylor series in λ about the origin, then its radius of convergence would be $2/\pi$.

- *Example 3*: Suppose that we need to solve the Fredholm integral equation

$$\psi(x) = x^4 + \tfrac{1}{10}\int_0^1 \mathrm{e}^{xt}\,\psi(t)\,\mathrm{d}t. \tag{2.6}$$

Since the kernel $K(x,t) = \mathrm{e}^{xt}$ is not separable, the methods of Chap. 1 do not apply.

In an attempt to use one of the successive methods, we easily compute

$$K_2(x,t) = \frac{\mathrm{e}^{x+t}-1}{x+t}$$

but then quickly realize that the integration required to produce the next iterate $K_3(x,t)$ is unmanageable.

Next, in an attempt to use the Method of Successive Approximation, we choose $\psi_0(x) = x^4$ and then use integration by parts several times to compute $\psi_1(x)$ as prescribed, but then we quickly find that the integration required to produce $\psi_2(x)$ is unmanageable.

In situations like this, *it may be advantageous to approximate the kernel* in our attempt to approximate the solution. Recalling that

$$\mathrm{e}^{xt} = \sum_{m=0}^{\infty} \frac{1}{m!}\,x^m t^m,$$

we truncate this series and consider the Fredholm integral equation

$$\phi(x) = x^4 + \tfrac{1}{10} \int_0^1 K(x,t)\,\phi(t)\,\mathrm{d}t \tag{2.7}$$

with the separable kernel

$$K(x,t) = 1 + xt + \frac{x^2t^2}{2!} + \frac{x^3t^3}{3!} + \frac{x^4t^4}{4!}$$

defined on the square $Q(0,1)$. All five eigenvalues of the kernel are real and positive, the smallest of which is $\lambda_1 \approx 0.739241$. Since $\frac{1}{10}$ is therefore a regular value, the integral equation has a unique solution by the First Fredholm Theorem.

A direct manual computation of the resolvent kernel is impractical since each of the iterated kernels $K_m(x,t)$ consists of 25 terms of the form $x^i y^j$. Instead, we prefer to determine a suitable approximation to the solution by using the Method of Successive Approximation. Since

$$\|K\|_2 = \frac{\sqrt{841,676,993}}{15120\sqrt{2}},$$

the Theorem of Successive Approximation guarantees that the resolvent series converges if $|\lambda| \le 1/\|K\|_2 \approx 0.737045$, a value that is remarkably close to λ_1. (The Theorem of Successive Substitution guarantees that the resolvent series converges if $|\lambda| \le 1/M(b-a) = \frac{24}{65} \approx 0.369231$.)

With the uniform convergence of the sequence $\{\phi_n(x)\}$ to the solution $\phi(x)$ now guaranteed, we can begin to compute it. Since $\lambda = \frac{1}{10}$ is small, we choose $\phi_0(x) = x^4$. In order to demonstrate the efficacy of the method to produce a convergent sequence, we list the first six approximants in the form

$$\phi_n(x) = x^4 + \tfrac{1}{10}\,\tau_n(x).$$

The coefficients of the polynomials $\tau_n(x)$ have been computed accurately here to six decimal places but not rounded up or down:

$$\begin{aligned}
\tau_1(x) &= 0.200000 + 0.166666x + 0.071428x^2 + 0.020833x^3 + 0.004629x^4\\
\tau_2(x) &= 0.231327 + 0.184501x + 0.077766x^2 + 0.022479x^3 + 0.004967x^4\\
\tau_3(x) &= 0.235611 + 0.186859x + 0.078590x^2 + 0.022692x^3 + 0.005010x^4\\
\tau_4(x) &= 0.236191 + 0.187178x + 0.078701x^2 + 0.022720x^3 + 0.005016x^4\\
\tau_5(x) &= 0.236269 + 0.187221x + 0.078717x^2 + 0.022724x^3 + 0.005017x^4\\
\tau_6(x) &= 0.236280 + 0.187226x + 0.078719x^2 + 0.022725x^3 + 0.005017x^4
\end{aligned}$$

The exact solution is given by $\phi(x) = x^4 + \frac{1}{10}\,\tau(x)$, where

$$\tau(x) = 0.236282 + 0.187227\,x + 0.078719\,x^2 + 0.022725\,x^3 + 0.005017\,x^4.$$

The rate of convergence is quite satisfactory. Indeed, the coefficients of $\tau(x)$ differ from those of $\tau_6(x)$ by at most 0.000002!

The degree of accuracy on display here could have been predicted by the error estimate given above. Since $\lambda = \frac{1}{10}$, $\|f\|_2 = \frac{1}{3}$, $\|K\|_2 = 1.3567691$, and $\|\sqrt{\kappa_1}\|_\infty = 1.7851749$, the error in using $\phi_6(x)$ to approximate $\phi(x)$ is no more than $|\phi(x) - \phi_6(x)| \le \frac{1}{10}(0.0000043)$.

To grade the success in approximating the solution $\psi(x)$ of Eq. (2.6) with the solution $\phi(x)$ of Eq. (2.7), it is reasonable to estimate the norm $\|\psi - \phi\|_\infty$.

An application of the triangle inequality yields

$$|\psi(x) - \phi(x)| \le \tfrac{1}{10} \left| \int_0^1 \mathrm{e}^{xt} \left(\psi(t) - \phi(t)\right) \mathrm{d}t \right| + \tfrac{1}{10} \left| \int_0^1 \left(\mathrm{e}^{xt} - K(x,t)\right) \phi(t)\, \mathrm{d}t \right|.$$

For each fixed $x \in [0,1]$, we apply the Cauchy–Schwarz inequality to each of these integrals to obtain

$$|\psi(x) - \phi(x)| \le \tfrac{1}{10} \|\mathrm{e}^{xt}\|_2 \cdot \|\psi - \phi\|_2 + \tfrac{1}{10} \|\mathrm{e}^{xt} - K(x,t)\|_2 \cdot \|\phi\|_2.$$

On the square $Q(0,1)$, we have

$$\|\mathrm{e}^{xt}\|_2 < \tfrac{9}{5}, \quad \text{and} \quad \|\mathrm{e}^{xt} - K(x,t)\|_2 < \tfrac{1}{338}$$

independently of x, so that the last inequality becomes

$$|\psi(x) - \phi(x)| \le \tfrac{9}{50} \|\psi - \phi\|_2 + \tfrac{1}{3380} \|\phi\|_2.$$

Taking the supremum over the left side of this inequality and replacing $\|\psi - \phi\|_2$ with the greater norm $\|\psi - \phi\|_\infty$ gives

$$\|\psi - \phi\|_\infty \le \tfrac{9}{50} \|\psi - \phi\|_\infty + \tfrac{1}{3380} \|\phi\|_2.$$

Finally, since

$$\|\phi\|_2 \le \|\phi_6\|_2 + \|\phi - \phi_6\|_2 \le \|\phi_6\|_2 + \|\phi - \phi_6\|_\infty < \tfrac{3}{8},$$

we obtain the uniform estimate

$$\|\psi - \phi\|_\infty \le \tfrac{50}{41} \cdot \tfrac{1}{3380} \cdot \tfrac{3}{8} = \tfrac{15}{110864} < \tfrac{1}{7390}.$$

Also, given the error estimate for $\|\phi - \phi_6\|_\infty$ that was established above, we have

$$\|\psi - \phi_6\|_\infty \le \|\psi - \phi\|_\infty + \|\phi - \phi_6\|_\infty < \tfrac{1}{7366}.$$

Practically speaking, the graphs of $\psi(x)$ and $\phi_6(x)$ are "within the ink" of each other, since their values agree so closely everywhere within the interval $[0,1]$.

Section 2.3 Exercises

1. Solve the nonhomogeneous Fredholm integral equation

$$\phi(x) = ax + \lambda \int_0^1 xt\,\phi(t)\,\mathrm{d}t$$

in two ways.

(a) Use the techniques in Chap. 1 to show that the unique solution to this equation is given by

$$\phi(x) = \frac{ax}{1 - \frac{\lambda}{3}}$$

if $\lambda \neq 3$.

(b) Use the Method of Successive Approximation, with the initial choice $\phi_0(x) = ax$, to compute $\phi_n(x)$, for every $n \geq 1$. Show that the limit $\phi(t) = \lim_{n\to\infty} \phi_n(x)$ exists. The solution obtained here should agree with the solution obtained above within the circle of convergence.

2. Suppose that the kernel $K(x,t)$ is continuous on the square $Q(a,b)$.

(a) If $K(x,x) \equiv 0$ on $[a,b]$, is it true that $K_2(x,x) \equiv 0$ on $[a,b]$?
(b) If $K(x,t) = K(t,x)$, is it true that $K_m(x,t) = K_m(t,x)$?
(c) If $K(x,t) = -K(t,x)$, is it true that $K_m(x,t) = (-1)^m K_m(t,x)$?

3. Suppose that the kernels $K(x,t)$ and $L(x,t)$ are continuous on the square $Q(a,b)$. If $K_2(x,t) = L_2(x,t)$, is it true that $K(x,t) = L(x,t)$?
4. Suppose that the kernel $K(x,t)$ is continuous on the symmetric square $Q(-a,+a)$ and that

$$K(x,y) = K(-x,y) = K(x,-y) = K(-x,-y).$$

Does $K_m(x,t)$ have the same property for all $m \geq 2$?
5. Suppose that the kernel $K(x,t)$ is continuous on the square $Q(a,b)$ and that $K(x,t) \equiv 0$ for all $a \leq x \leq t \leq b$, i.e., the kernel vanishes on and above the diagonal of the square. Does $K_2(x,t)$ have the same property?
6. Define

$$K(x,t) = \begin{cases} t\,(1-x) & \text{if } t \leq x \\ x\,(1-t) & \text{if } x \leq t \end{cases}$$

on the square $Q(0,1)$. This kernel vanishes on the boundary of the square, and $K(x,x) = x(1-x)$ on its diagonal. Compute $K_2(x,t)$.
7. Define $K(x,t) = |x-t|$ on the square $Q(0,a)$. Compute $K_2(x,t)$.

8. By using its series representation, show that the resolvent satisfies the equations

$$R(x,t;\lambda) = K(x,t) + \lambda \int_a^b K(x,s)\, R(s,t;\lambda)\, \mathrm{d}s$$

and

$$R(x,t;\lambda) = K(x,t) + \lambda \int_a^b R(x,s;\lambda)\, K(s,t)\, \mathrm{d}s.$$

9. By using its series representation, show that the resolvent satisfies the integro-differential equation

$$\frac{\partial}{\partial \lambda} R(x,t;\lambda) = \int_a^b R(x,s;\lambda)\, R(s,t;\lambda)\, \mathrm{d}s.$$

10. Let $K(x,t)$ be a continuous kernel defined on $Q(a,b)$, and let

$$\|K\|_2 = \left(\int_a^b \int_a^b |K(x,t)|^2 \, \mathrm{d}x \, \mathrm{d}t \right)^{1/2}.$$

Assume that λ is small enough so that

$$\|R_\lambda\|_2 = \left(\int_a^b \int_a^b |R(x,t;\lambda)|^2 \, \mathrm{d}x \, \mathrm{d}t \right)^{1/2} < +\infty.$$

Estimate $\|R_\lambda\|_2$ in terms of $|\lambda|$ and $\|K\|_2$. For which values of λ is your estimate valid?

11. Consider the integral equation

$$\psi(x) = x + \tfrac{1}{2} \int_0^1 \cosh(xt)\, \psi(t)\, \mathrm{d}t$$

and assume that $\frac{1}{2}$ is not an eigenvalue of the kernel. Choose $\psi_0(x) = x$ and compute $\psi_1(x)$, $\psi_2(x)$, and $\psi_3(x)$. (Recall that $\cosh x = (\mathrm{e}^x + \mathrm{e}^{-x})/2$.) Next, approximate the kernel by

$$\cosh(xt) \approx 1 + \tfrac{1}{2} x^2 t^2 + \tfrac{1}{24} x^4 t^4$$

and consider the integral equation

$$\phi(x) = x + \tfrac{1}{2} \int_0^1 \left(1 + \tfrac{1}{2} x^2 t^2 + \tfrac{1}{24} x^4 t^4\right) \phi(t)\, \mathrm{d}t.$$

Choose $\phi_0(x) = x$, and compute $\phi_1(x)$, $\phi_2(x)$, and $\phi_3(x)$. Compare your results, and estimate the accuracy of your approximations by following the techniques in Example 3.

2.4 The Fredholm Theorems

In Sect. 1.3, we proved the Fredholm theorems for the integral equation

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t, \tag{2.8}$$

where λ is an arbitrary complex parameter and $K(x,t)$ is a separable kernel. Then, in Sect. 2.3, we showed that the Method of Successive Approximations can be successfully applied to solve this equation if $K(x,t)$ is a general kernel, provided that $|\lambda| < 1/\|K\|_2$. In this section, we combine the techniques of those two sections in order to establish the Fredholm theorems if $K(x,t)$ is a general kernel and λ is arbitrary.

An ingenious idea due to Schmidt enables the discussion. His idea is to show that the solution $\phi(x)$ to Eq. (2.8) satisfies *two* integral equations, one whose kernel is small and another one whose kernel is separable. The techniques delineated in Sects. 1.3 and 2.3 can then be applied to these equations.

Some assumptions are in order. The complex parameter λ is assumed at first to belong to the closed disk $\Delta_\rho = \{\lambda : |\lambda| \le \rho\}$, where ρ is arbitrarily large but fixed. As usual, the free term $f(x)$ is assumed to be complex-valued and continuous on the interval $[a,b]$, and the kernel $K(x,t)$ is assumed to be complex-valued and continuous on the square $Q(a,b)$.

The discussion begins with a decomposition of the kernel $K(x,t)$.

By virtue of the well-known Weierstrass approximation theorem, the kernel $K(x,t)$ can be decomposed into a sum of two complex-valued and continuous kernels as

$$K(x,t) = K_{\mathrm{sep}}(x,t) + K_\varepsilon(x,t).$$

The kernel $K_{\mathrm{sep}}(x,t)$ can be chosen to be a separable polynomial in the variables x and t in the form

$$K_{\mathrm{sep}}(x,t) = \sum_{i=1}^{n} a_i(x)\,\overline{b_i(t)},$$

where each $a_i(x)$ and $b_i(t)$ is complex-valued and continuous on $[a,b]$, and each of the sets $\{a_i(x)\}$ and $\{b_i(t)\}$ is linearly independent on the interval $[a,b]$. The kernel $K_\varepsilon(x,t)$ can be chosen so that its norm

$$\|K_\varepsilon\|_2 = \left(\int_a^b \int_a^b |K_\varepsilon(x,t)|^2\,\mathrm{d}x\,\mathrm{d}t \right)^{1/2}$$

is arbitrarily small. In particular, if we require that $\|K_\varepsilon\|_2 < \varepsilon$, we choose $\rho = 1/\varepsilon$, so that if $|\lambda| \le \rho$, then $|\lambda| < 1/\|K_\varepsilon\|_2$.

Given this decomposition of the kernel, we may write Eq. (2.8) in the form

$$\phi(x) = f(x) + \lambda \int_a^b K_{\mathrm{sep}}(x,t)\,\phi(t)\,\mathrm{d}t + \lambda \int_a^b K_\varepsilon(x,t)\,\phi(t)\,\mathrm{d}t.$$

Equivalently, we have

$$\phi(x) = F(x;\lambda) + \lambda \int_a^b K_\varepsilon(x,t)\,\phi(t)\,\mathrm{d}t, \tag{2.9}$$

which is the first of the two integral equations satisfied by $\phi(x)$, where

$$F(x;\lambda) = f(x) + \lambda \int_a^b K_{\mathrm{sep}}(x,t)\,\phi(t)\,\mathrm{d}t.$$

The continuity of $F(x;\lambda)$ on the interval $[a,b]$ is due to the integrability of $\phi(t)$. At least for $|\lambda| < 1/\|K\|_2$, the solution $\phi(x)$ that exists by virtue of the Theorem of Successive Approximation is continuous on the interval $[a,b]$.

Now consider the equation

$$\psi(x) = g(x) + \lambda \int_a^b K_\varepsilon(x,t)\,\psi(t)\,\mathrm{d}t,$$

where $g(x)$ is assumed to be complex-valued and continuous on the interval $[a,b]$. As a consequence of the Theorem of Successive Approximation, the solution $\psi(x)$ to this equation may be expressed in the form

$$\psi(x) = g(x) + \lambda \int_a^b R_\varepsilon(x,t;\lambda)\,g(t)\,\mathrm{d}t,$$

if $|\lambda| < 1/\|K_\varepsilon\|_2$. In particular, this conclusion holds for all values of $\lambda \in \Delta_\rho$, since we chose $K_\varepsilon(x,t)$ so that $\rho < 1/\|K_\varepsilon\|_2$. As a consequence of this choice, the resolvent series

$$R_\varepsilon(x,t;\lambda) = \sum_{m=1}^{\infty} \lambda^{m-1} K_{\varepsilon m}(x,t), \tag{2.10}$$

generated from the iterates of $K_\varepsilon(x,t)$, is an analytic function of λ in Δ_ρ and is a continuous function of x and t on the square $Q(a,b)$ for each $\lambda \in \Delta_\rho$.

Whenever $F(x;\lambda)$ is continuous, the solution to Eq. (2.9) can be displayed as

$$\phi(x) = F(x;\lambda) + \lambda \int_a^b R_\varepsilon(x,t;\lambda)\,F(t;\lambda)\,\mathrm{d}t.$$

After replacing $F(x;\lambda)$ by its definition in this equation, we obtain

$$\phi(x) = f_\varepsilon(x;\lambda) + \lambda \int_a^b G_\varepsilon(x,t;\lambda)\,\phi(t)\,\mathrm{d}t, \tag{2.11}$$

where

$$f_\varepsilon(x;\lambda) = f(x) + \lambda \int_a^b R_\varepsilon(x,t;\lambda)\,f(t)\,\mathrm{d}t \tag{2.12}$$

and

$$G_\varepsilon(x,t;\lambda) = K_{\text{sep}}(x,t) + \lambda \int_a^b R_\varepsilon(x,u;\lambda)\, K_{\text{sep}}(u,t)\, \mathrm{d}u. \tag{2.13}$$

Equation (2.11) is the second integral equation that is satisfied by the solution $\phi(x)$ to Eq. (2.8). The free term $f_\varepsilon(x;\lambda)$ is an analytic function of λ in Δ_ρ, is complex-valued and continuous on the interval $[a,b]$, and exhibits an adjustment to the original free term $f(x)$ that is based on the choice of $K_\varepsilon(x,t)$. The kernel $G_\varepsilon(x,t;\lambda)$ is an analytic function of λ on Δ_ρ and is complex-valued and continuous on the square $Q(a,b)$. It reflects an adjustment to the original kernel $K(x,t)$ that is based upon its decomposition.

It is important to note that $G_\varepsilon(x,t;\lambda)$ is separable. To see this, observe that

$$\begin{aligned}
\int_a^b R_\varepsilon(x,u;\lambda)\, K_{\text{sep}}(u,t)\, \mathrm{d}u &= \int_a^b R_\varepsilon(x,u;\lambda) \left(\sum_{i=1}^n a_i(u)\, \overline{b_i(t)} \right) \mathrm{d}u \\
&= \sum_{i=1}^n \left(\int_a^b R_\varepsilon(x,u;\lambda)\, a_i(u)\, \mathrm{d}u \right) \overline{b_i(t)} \\
&= \sum_{i=1}^n A_{\varepsilon i}(x;\lambda)\, \overline{b_i(t)},
\end{aligned}$$

where

$$\begin{aligned}
A_{\varepsilon i}(x;\lambda) &= \int_a^b R_\varepsilon(x,u;\lambda)\, a_i(u)\, \mathrm{d}u \\
&= \int_a^b \left(\sum_{m=1}^\infty \lambda^{m-1} K_{\varepsilon m}(x,u) \right) a_i(u)\, \mathrm{d}u \\
&= \sum_{m=1}^\infty \lambda^{m-1} \left(\int_a^b K_{\varepsilon m}(x,u)\, a_i(u)\, \mathrm{d}u \right).
\end{aligned} \tag{2.14}$$

Term-by-term integration is permitted here, since the resolvent series converges absolutely and uniformly on the interval $[a,b]$ for every $\lambda \in \Delta_\rho$. Each $A_{\varepsilon i}(x;\lambda)$ is an analytic function of λ on Δ_ρ and is continuous on the interval $[a,b]$. With these notational conventions, the kernel $G_\varepsilon(x,t;\lambda)$ can be explicitly displayed in the separable form

$$G_\varepsilon(x,t;\lambda) = \sum_{i=1}^n \left[a_i(x) + \lambda\, A_{\varepsilon i}(x;\lambda) \right] \overline{b_i(t)}. \tag{2.15}$$

Equation (2.11) can now be treated with the techniques that were delineated in Sect. 1.3 for integral equations with separable kernels, although some differences will arise due to the fact that its kernel depends upon λ. The analysis of Eq. (2.11) begins by replacing $G_\varepsilon(x,t;\lambda)$ in with its representation (2.15) to obtain

$$\phi(x) = f_\varepsilon(x;\lambda) + \lambda \sum_{i=1}^{n} c_i(\lambda)\,[a_i(x) + \lambda A_{\varepsilon i}(x;\lambda)]\,, \tag{2.16}$$

where we have set

$$c_i(\lambda) = \int_a^b \phi(t)\,\overline{b_i(t)}\,\mathrm{d}t.$$

Every solution to Eq. (2.11) assumes this form. It remains to determine the coefficients $c_i(\lambda)$. If we replace x by t, change the index of summation from i to j, multiply both sides of Eq. (2.16) by $\overline{b_i(t)}$, and then integrate the resulting expression from a to b, then we obtain the system of linear equations

$$c_i(\lambda) = f_i(\lambda) + \lambda \sum_{j=1}^{n} c_j(\lambda)\,a_{ij}(\lambda) \tag{2.17}$$

for $i = 1,\ldots,n$, where we have set

$$f_i(\lambda) = \int_a^b f_\varepsilon(t;\lambda)\,\overline{b_i(t)}\,\mathrm{d}t$$

and

$$a_{ij}(\lambda) = \int_a^b [a_j(t) + \lambda A_{\varepsilon j}(t;\lambda)]\,\overline{b_i(t)}\,\mathrm{d}t. \tag{2.18}$$

Since the definitions of $f_i(\lambda)$ and $a_{ij}(\lambda)$ depend upon the resolvent kernel $R_\varepsilon(x,t;\lambda)$, they both are analytic functions of λ in Δ_ρ.

The linear system (2.17) can also be written in matrix form as

$$(\mathbf{I} - \lambda\,\mathbf{A}(\lambda))\,\mathbf{c}(\lambda) = \mathbf{f}(\lambda). \tag{2.19}$$

The Fredholm determinant $D_\rho(\lambda) = \det(\mathbf{I} - \lambda\,\mathbf{A}(\lambda))$ is an analytic function of λ in the closed disk Δ_ρ. An analytic function defined on an open neighborhood of a closed disk D in the complex plane that vanishes on an infinite set with a limit point in D vanishes identically on the neighborhood of D. Since $D_\rho(0) = 1$, $D_\rho(\lambda)$ can have only a finite number of zeroes in Δ_ρ. In Sect. 1.3, the matrix $\mathbf{A}$, corresponding to the matrix $\mathbf{A}(\lambda)$ here, was a matrix of constants, and the Fredholm determinant $D(\lambda) = \det(\mathbf{I} - \lambda\,\mathbf{A})$ was a polynomial of degree n with exactly n zeroes. However, since $D_\rho(\lambda)$ is not necessarily a polynomial, it may have more or less than n zeroes, but in any case, the number of zeroes is finite.

The values of $\lambda \in \Delta_\rho$ for which $D_\rho(\lambda) \neq 0$ are called *regular* values of the kernel $K(x,t)$. If λ is a regular value, then the unique solution $\phi(x)$ to the inhomogeneous integral equation (2.8) can ultimately be expressed as an integral with a resolvent kernel as stated below in the First Fredholm Theorem.

On the other hand, the values of λ for which $D_\rho(\lambda) = 0$ are called *eigenvalues of the kernel* $K(x,t)$. If λ is an eigenvalue, then the homogeneous case of the integral

equation (2.8) will have nontrivial solutions called *eigenfunctions of the kernel*; the inhomogeneous case may or may not have solutions, depending upon additional considerations.

The analysis of the linear system (2.19) proceeds in remarkedly different ways, depending upon the value of the Fredholm determinant $D_\rho(\lambda)$. The linear system, and hence the integral equation (2.8), will have exactly one solution if the determinant does not vanish. Otherwise, the linear system will have either an infinite number of solutions or no solutions at all. Both possibilities may occur, and each of them has implications for our investigation.

We consider two cases:

Case I: $D_\rho(\lambda) = \det(\mathbf{I} - \lambda\,\mathbf{A}(\lambda)) \neq 0$.

In this case, the linear system has the unique solution

$$\begin{aligned}\mathbf{c}(\lambda) &= (\mathbf{I} - \lambda\,\mathbf{A}(\lambda))^{-1}\mathbf{f}(\lambda) \\ &= \frac{1}{D_\rho(\lambda)}\,\mathrm{adj}(\mathbf{I} - \lambda\,\mathbf{A}(\lambda))\,\mathbf{f}(\lambda),\end{aligned} \tag{2.20}$$

where $\mathrm{adj}(\mathbf{I} - \lambda\,\mathbf{A}(\lambda)) = (D_{ji}(\lambda))$ is the transpose of the matrix of cofactors from $\mathbf{I} - \lambda\,\mathbf{A}(\lambda)$. Each coefficient $c_i(\lambda)$ can be represented in the form

$$c_i(\lambda) = \frac{1}{D_\rho(\lambda)} \sum_{j=1}^{n} D_{ji}(\lambda)\, f_j(\lambda).$$

Note that each $D_{ji}(\lambda)$ is an analytic function of λ in the closed disk Δ_ρ.

We consider two possibilities:

1. If $\mathbf{f}(\lambda) = \mathbf{0}$, then $\mathbf{c}(\lambda) = \mathbf{0}$, i.e., $c_i(\lambda) = 0$ for all $i = 1, \ldots, n$. Since every solution to Eq. (2.11) has the general form (2.16), it follows that $\phi(x) = f_\varepsilon(x;\lambda)$. In particular, if $f(x) \equiv 0$, then $f_\varepsilon(x;\lambda) \equiv 0$ on $[a,b]$, so that the unique solution to Eq. (2.11) is $\phi(x) \equiv 0$. It is possible for $\mathbf{f}(\lambda) = \mathbf{0}$ with $f_\varepsilon(x;\lambda) \neq 0$ on $[a,b]$. For instance, if each $b_i(t)$ happens to be orthogonal to $f(t)$ and the iterates of $K_\varepsilon(t,u)$, then $\mathbf{f}(\lambda) = \mathbf{0}$.
2. If $\mathbf{f}(\lambda) \neq \mathbf{0}$, then $\mathbf{c}(\lambda) \neq \mathbf{0}$, i.e., $c_i(\lambda) \neq 0$ for at least one subscript i. After substituting these values for $c_i(\lambda)$ into Eq. (2.16), we obtain

$$\begin{aligned}\phi(x) &= f_\varepsilon(x;\lambda) + \lambda \sum_{i=1}^{n} \left(\frac{1}{D_\rho(\lambda)} \sum_{j=1}^{n} D_{ji}(\lambda)\, f_j(\lambda) \right) [a_i(x) + \lambda\, A_{\varepsilon i}(x;\lambda)] \\ &= f_\varepsilon(x;\lambda) + \lambda \int_a^b S_\varepsilon(x,t;\lambda)\, f_\varepsilon(t;\lambda)\, \mathrm{d}t,\end{aligned}$$

where we have set

$$S_\varepsilon(x,t;\lambda) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} D_{ji}(\lambda)\, [a_i(x) + \lambda\, A_{\varepsilon i}(x;\lambda)]\, \overline{b_j(t)}}{D_\rho(\lambda)}. \tag{2.21}$$

The kernel $S_\varepsilon(x,t;\lambda)$ is a meromorphic function of λ in the closed disk Δ_ρ, in addition to being a separable function of x and t. Since $S_\varepsilon(x,t;0) = K_{\text{sep}}(x,t)$, $S_\varepsilon(x,t;\lambda)$ represents an adjustment to $K_{\text{sep}}(x,t)$ based upon the decomposition of $K(x,t)$ and can also be written as a negative quotient of determinants.

After substituting $f_\varepsilon(x;\lambda)$ into this last representation for $\phi(x)$, we finally obtain a representation for the solution to Eq. (2.11), and therefore also Eq. (2.8), in the compact form

$$\phi(x) = f(x) + \lambda \int_a^b U_\varepsilon(x,t;\lambda)\, f(t)\, \mathrm{d}t,$$

where

$$U_\varepsilon(x,t;\lambda) = R_\varepsilon(x,t;\lambda) + S_\varepsilon(x,t;\lambda) + \lambda \int_a^b S_\varepsilon(x,s;\lambda)\, R_\varepsilon(s,t;\lambda)\, \mathrm{d}s. \tag{2.22}$$

The resolvent $R_\varepsilon(x,t;\lambda)$ is relatively small, since it was obtained by iterating $K_\varepsilon(x,t)$, and the integral term is also small, since $R_\varepsilon(s,t;\lambda)$ appears in the integrand. Hence, $S_\varepsilon(x,t;\lambda)$ is the main component of $U_\varepsilon(x,t;\lambda)$.

The kernel $U_\varepsilon(x,t;\lambda)$ is a meromorphic function of λ in the closed disk Δ_ρ. Actually, $U_\varepsilon(x,t;\lambda)$ is the restriction of a function that is meromorphic in the full λ-plane. For if $\tilde{\rho}$ is arbitrary with $0 < \rho < \tilde{\rho}$, then we can construct another resolvent kernel $\tilde{U}_\varepsilon(x,t;\lambda)$ that is meromorphic in $\Delta_{\tilde{\rho}}$, so that the solution to Eq. (2.11) can also be represented in the form

$$\phi(x) = f(x) + \lambda \int_a^b \tilde{U}_\varepsilon(x,t;\lambda)\, f(t)\, \mathrm{d}t.$$

However, since $\phi(x)$ is unique, $U_\varepsilon(x,t;\lambda) = \tilde{U}_\varepsilon(x,t;\lambda)$ on a neighborhood of the origin. Thus, $\tilde{U}_\varepsilon(x,t;\lambda)$ is a meromorphic extension of $U_\varepsilon(x,t;\lambda)$. Since $\tilde{\rho}$ was assumed to be arbitrary, it follows $U_\varepsilon(x,t;\lambda)$ may be extended to the entire λ-plane, and we let $R(x,t;\lambda)$ denote this unique extension.

The discussion above serves to establish the following result:

Theorem 2.4.1 (The First Fredholm Theorem). *Let λ be a complex parameter, $f(x)$ be a complex-valued continuous function defined on the interval $[a,b]$, and $K(x,t)$ be a complex-valued continuous kernel defined on the region $Q(a,b)$.*

Then the unique solution $\phi(x)$ to the Fredholm integral equation of the second kind

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\, \phi(t)\, \mathrm{d}t$$

has the representation

$$\phi(x) = f(x) + \lambda \int_a^b R(x,t;\lambda)\, f(t)\, \mathrm{d}t,$$

for every regular value of λ, where the unique resolvent kernel $R(x,t;\lambda)$ is a meromorphic function of λ in the complex plane $\mathbb{C}$.

Note that if λ is a regular value of the kernel, then the homogeneous integral equation has only the trivial solution $\phi(x) \equiv 0$ on the interval $[a,b]$. For if $f(x) \equiv 0$ on $[a,b]$, then $f_\varepsilon(x;\lambda) \equiv 0$, so that $\mathbf{f}(\lambda) = \mathbf{0}$, implying in turn that $\mathbf{c}(\lambda) = \mathbf{0}$. The general form (2.16) of the solution reduces to $\phi(x) \equiv 0$.

In the discussion above, it was also established that the disk Δ_ρ, ρ arbitrary, could contain only a finite number of eigenvalues of the kernel $K(x,t)$. If the set of eigenvalues of $K(x,t)$ had a finite limit point, then there would exist some value of ρ for which Δ_ρ contained an infinite number of eigenvalues. The next result follows from this observation.

Theorem 2.4.2 (The Fourth Fredholm Theorem). *Let $K(x,t)$ be a complex-valued continuous kernel defined on the square $Q(a,b)$. Let Λ_K denote the set of eigenvalues of the kernel $K(x,t)$, that is, the set of values λ for which the homogeneous Fredholm integral equation*

$$\phi(x) = \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t$$

has a nontrivial solution. Then Λ_K is at most countable, and it cannot have a finite limit point.

Case II: $D_\rho(\lambda) = \det(\mathbf{I} - \lambda\,\mathbf{A}(\lambda)) = 0$.

In this case, there are a finite number of values $\lambda \in \Delta_\rho$ for which $D_\rho(\lambda) = 0$. We consider two possibilities:

1. If $\mathbf{f}(\lambda) = \mathbf{0}$, then the linear system

$$(\mathbf{I} - \lambda\,\mathbf{A}(\lambda))\,\mathbf{c}(\lambda) = \mathbf{0} \tag{2.23}$$

has a certain number $p(\lambda)$ of nonzero linearly independent vector solutions $\mathbf{c}^{(j)}(\lambda)$, $j = 1,\ldots,p(\lambda)$, that can be written in the form

$$\mathbf{c}^{(j)}(\lambda) = \begin{pmatrix} c_1^{(j)}(\lambda) \\ \vdots \\ c_n^{(j)}(\lambda) \end{pmatrix}.$$

Insert these values of $c_i^{(j)}(\lambda)$ into Eq. (2.16) to obtain the solution to Eq. (2.8). If $f(x) \equiv 0$ on $[a,b]$, then $f_\varepsilon(x;\lambda) \equiv 0$. For $j = 1,\ldots,p(\lambda)$, the general form of the solution to the homogeneous counterpart to the integral equation (2.8) reduces to

$$\begin{aligned} \phi_j^{(e)}(x;\lambda) &= \sum_{i=1}^{n} c_i^{(j)}(\lambda)\,[a_i(x) + \lambda\,A_{\varepsilon i}(x;\lambda)] \\ &= \sum_{i=1}^{n} c_i^{(j)}(\lambda)\left[a_i(x) + \lambda \int_a^b R_\varepsilon(x,t;\lambda)\,a_i(u)\,\mathrm{d}u\right], \end{aligned} \tag{2.24}$$

where the superscript (e) signifies that $\phi_j^{(e)}(x;\lambda)$ is an eigenfunction of the kernel $K(x,t)$. The span of the set of all eigenfunctions corresponding to the eigenvalue λ constitutes an eigenspace, whose dimension is $p(\lambda)$. The general solution to the homogeneous integral equation corresponding to a given eigenvalue λ assumes the form

$$\phi^{(h)}(x;\lambda) = \sum_{j=1}^{p(\lambda)} \alpha_j \phi_j^{(e)}(x;\lambda),$$

where the α_j are arbitrary constants. The superscript (h) signifies that the solution $\phi^{(h)}(x;\lambda)$ is the most general solution to the homogeneous equation corresponding to the eigenvalue λ.

The number $p(\lambda)$ of linearly independent eigenfunctions corresponding to a given eigenvalue λ is called the *geometric multiplicity of the eigenvalue*. The geometric multiplicity of the eigenvalue is less than or equal to the algebraic multiplicity of the eigenvalue, i.e., if λ is an $m(\lambda)$-fold root of the equation $D_\rho(\lambda) = 0$, then $1 \leq p(\lambda) \leq m(\lambda)$.

When λ is an eigenvalue of a separable kernel, the rationale for considering the homogeneous adjoint integral equation while conducting an analysis of the inhomogeneous integral equation (2.8) was firmly established in Sect. 1.3. Here, we examine this relationship anew when λ is an eigenvalue of a general kernel.

Consider the homogeneous adjoint integral equation

$$\psi(x) = \overline{\lambda} \int_a^b \overline{K(t,x)}\, \psi(t)\, \mathrm{d}t \tag{2.25}$$

where $K(x,t)$ is complex-valued and continuous on the square $Q(a,b)$. If we conjugate the decomposition of the kernel and transpose variables, then we obtain

$$\overline{K(t,x)} = \overline{K_{\mathrm{sep}}(t,x)} + \overline{K_\varepsilon(t,x)}.$$

If we now substitute this decomposition into the homogeneous adjoint equation and rearrange terms, then we have

$$\omega(x) = \psi(x) - \overline{\lambda} \int_a^b \overline{K_\varepsilon(t,x)}\, \psi(t)\, \mathrm{d}t = \overline{\lambda} \int_a^b \overline{K_{\mathrm{sep}}(t,x)}\, \psi(t)\, \mathrm{d}t. \tag{2.26}$$

It is clear from this representation that if $\psi(x) \equiv 0$, then $\omega(x) \equiv 0$ on $[a,b]$. Since the left half of this equation is an inhomogeneous Fredholm equation with a small kernel, the Theorem of Successive Approximation can be applied to it, thereby obtaining a representation for $\psi(x)$ in the form

$$\psi(x) = \omega(x) + \overline{\lambda} \int_a^b \overline{R_\varepsilon(t,x;\lambda)}\, \omega(t)\, \mathrm{d}t, \tag{2.27}$$

where the resolvent kernel $\overline{R_\varepsilon(t,x;\lambda)}$ is the complex conjugate of the resolvent kernel $R_\varepsilon(x,t;\lambda)$ constructed from the iterations of the kernel $K_\varepsilon(x,t)$. It is clear from this representation that if $\omega(x) \equiv 0$, then $\psi(x) \equiv 0$ on $[a,b]$. Thus, $\omega(x)$ and $\psi(x)$ vanish only simultaneously. If we substitute the right half of the representation (2.26) for $\omega(x)$ into the last equation, we obtain

$$\psi(x) = \overline{\lambda} \int_a^b \overline{G_\varepsilon(t,x;\lambda)}\, \psi(t)\, \mathrm{d}t, \tag{2.28}$$

where the kernel $\overline{G_\varepsilon(t,x;\lambda)}$ is the complex conjugate of the kernel $G_\varepsilon(x,t;\lambda)$ that appears in Eq. (2.11). It is explicitly given in the integral form (2.13) or in the equivalent separable form (2.15), from which we conclude that

$$\overline{G_\varepsilon(t,x;\lambda)} = \sum_{i=1}^{n} \overline{[a_i(t) + \lambda A_{\varepsilon i}(t;\lambda)]}\, b_i(x). \tag{2.29}$$

It follows that the solutions to Eq. (2.28) have the representation

$$\psi_j^{(e)}(x;\lambda) = \overline{\lambda} \sum_{i=1}^{n} d_i^{(j)}(\lambda)\, b_i(x),$$

for $j = 1, \ldots, q(\lambda)$, where $q(\lambda)$ is the geometric multiplicity of λ and we have set

$$d_i^{(j)}(\lambda) = \int_a^b \psi(t)\, \overline{[a_i(t) + \lambda A_{\varepsilon i}(t;\lambda)]}\, \mathrm{d}t.$$

We conclude that $\overline{\lambda}$ is an eigenvalue of the kernel $\overline{K(t,x)}$ and that all of its $q(\lambda)$ linearly independent solutions are as given. Since Eq. (2.28) and the equation

$$\phi(x) = \lambda \int_a^b G_\varepsilon(x,t;\lambda)\, \phi(t)\, \mathrm{d}t$$

have conjugate kernels, they must have the same number of linearly independent eigenfunctions, i.e., $p(\lambda) = q(\lambda)$.

The discussion above serves to establish the following result:

Theorem 2.4.3 (The Second Fredholm Theorem). *Let $K(x,t)$ be a complex-valued continuous kernel defined on the square $Q(a,b)$. If λ is an eigenvalue of the kernel $K(x,t)$, then $\overline{\lambda}$ is an eigenvalue of the adjoint kernel $\overline{K(t,x)}$. The number of linearly independent eigenfunctions of the homogeneous equation*

$$\phi(x) = \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t$$

is equal to the number of linearly independent eigenfunctions of the homogeneous adjoint equation

$$\psi(x) = \overline{\lambda} \int_a^b \overline{K(t,x)}\,\psi(t)\,\mathrm{d}t.$$

2. If $\mathbf{f}(\lambda) \neq \mathbf{0}$, then neither of the free terms $f_\varepsilon(x;\lambda)$ and $f(x)$ can vanish identically on the interval $[a,b]$, so that the equations

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t$$

and

$$\phi(x) = f_\varepsilon(x;\lambda) + \lambda \int_a^b G_\varepsilon(x,t;\lambda)\,\phi(t)\,\mathrm{d}t$$

are simultaneously inhomogeneous and are simultaneously satisfied by $\phi(x)$. Since $G_\varepsilon(x,t;\lambda)$ is separable, the Third Fredholm Theorem for integral equations with separable kernels that was proven in Sect. 1.3 can be invoked to conclude that the last equation has solutions if and only if $f_\varepsilon(x;\lambda)$ is orthogonal to all solutions of the homogeneous adjoint equation

$$\psi(x) = \overline{\lambda} \int_a^b \overline{G_\varepsilon(t,x;\lambda)}\,\psi(t)\,\mathrm{d}t.$$

If $\omega(x)$ denotes one of these solutions, then, given the representation (2.12), we have

$$\begin{aligned}
\int_a^b f_\varepsilon(t;\lambda)\,\overline{\omega(t)}\,\mathrm{d}t &= \int_a^b \left(f(t) + \lambda \int_a^b R_\varepsilon(t,s;\lambda)\,f(s)\,\mathrm{d}s \right) \overline{\omega(t)}\,\mathrm{d}t \\
&= \int_a^b f(t)\,\overline{\omega(t)}\,\mathrm{d}t + \lambda \int_a^b \int_a^b R_\varepsilon(t,s;\lambda)\,f(s)\,\overline{\omega(t)}\,\mathrm{d}s\,\mathrm{d}t \\
&= \int_a^b f(t) \left[\overline{\omega(t)} + \lambda \int_a^b R_\varepsilon(s,t;\lambda)\,\overline{\omega(s)}\,\mathrm{d}s \right] \mathrm{d}t \\
&= \int_a^b f(t) \overline{\left[\omega(t) + \overline{\lambda} \int_a^b \overline{R_\varepsilon(s,t;\lambda)}\,\omega(s)\,\mathrm{d}s \right]}\,\mathrm{d}t \\
&= \int_a^b f(t)\,\overline{\psi(t)}\,\mathrm{d}t.
\end{aligned}$$

The last equality follows from the representation (2.27). We conclude from these equations that $f_\varepsilon(x;\lambda)$ is orthogonal to $\omega(x)$ if and only if $f(x)$ is orthogonal to any solution $\psi(x)$ of the homogeneous adjoint equation

$$\psi(x) = \overline{\lambda} \int_a^b \overline{K(t,x)}\, \psi(t)\, \mathrm{d}t.$$

The discussion above serves to establish the following result:

Theorem 2.4.4 (The Third Fredholm Theorem). *Let λ be a complex parameter, $f(x)$ be a complex-valued continuous function defined on the interval $[a,b]$, and $K(x,t)$ be a complex-valued continuous kernel defined on the square $Q(a,b)$. If λ is an eigenvalue of the kernel $K(x,t)$, then the inhomogeneous Fredholm integral equation of the second kind*

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\, \phi(t)\, \mathrm{d}t$$

will have solutions if and only if the free term $f(x)$ is orthogonal to all of the eigenfunctions of the homogeneous adjoint equation

$$\psi(x) = \overline{\lambda} \int_a^b \overline{K(t,x)}\, \psi(t)\, \mathrm{d}t.$$

The Fredholm Alternative Theorem for integral equations with separable kernels was stated in Sect. 1.3. The statement of this theorem for general kernels is identical, except for the fact that the word "separable" is deleted.

Although the Fredholm theorems give the conditions under which solutions to Fredholm integral equations of the second kind exist, they do not specify how to construct those solutions. In the next section, we provide a recursive procedure for constructing the resolvent kernel when λ is regular.

Illustrative Examples

- *Example 1: The Lalesco–Picard Equation*. This example highlights the importance of the Fourth Fredholm Theorem.

  Consider the homogeneous integral equation

  $$\phi(x) = \lambda \int_{-b}^{+b} \mathrm{e}^{-|x-t|}\, \phi(t)\, \mathrm{d}t,$$

  where $0 < \lambda < +\infty$ and $0 < b \leq +\infty$. The presence of the absolute value in the integrand requires the equation to be expanded in the equivalent form

  $$\phi(x) = \lambda \left(\mathrm{e}^{-x} \int_{-b}^{x} \mathrm{e}^{t}\, \phi(t)\, \mathrm{d}t + \mathrm{e}^{x} \int_{x}^{+b} \mathrm{e}^{-t}\, \phi(t)\, \mathrm{d}t \right).$$

  After differentiation and simplification, we obtain the second-order linear ordinary differential equation

$$\phi''(x) + (2\lambda - 1)\,\phi(x) = 0.$$

Any solution to the Lalesco–Picard equation must necessarily be a solution to this differential equation, which has exactly two linearly independent solutions whose form depends upon the value of λ. These solutions are

$$\phi(x) = \begin{cases} A\sinh(\sqrt{1-2\lambda}) + B\cosh(\sqrt{1-2\lambda}), & \text{if } 0 < \lambda < \frac{1}{2} \\ A + Bx, & \text{if } \lambda = \frac{1}{2} \\ A\sin(\sqrt{2\lambda-1}) + B\cos(\sqrt{2\lambda-1}), & \text{if } \frac{1}{2} < \lambda < +\infty \end{cases}$$

It remains to determine, by process of elimination, exactly which of these six solutions to the differential equation are eigenfunctions of the kernel.

We consider three cases:

Case I: $0 < \lambda < \frac{1}{2}$.

In this case, we set $\mu = \sqrt{1-2\lambda}$ or $\lambda = (1-\mu^2)/2$, where $0 < \mu < 1$.

After substituting $\phi(x) = \sinh(\mu x)$ into the expanded integral equation, we obtain

$$\sinh(\mu x) = \sinh(\mu x) + \lambda\,\mathrm{e}^{-b}\sinh x\left(-\frac{\mathrm{e}^{b\mu}}{1-\mu} + \frac{\mathrm{e}^{-b\mu}}{\mu+1}\right);$$

after substituting $\phi(x) = \cosh(\mu x)$, we obtain

$$\cosh(\mu x) = \cosh(\mu x) - \lambda\,\mathrm{e}^{-b}\cosh x\left(\frac{\mathrm{e}^{b\mu}}{1-\mu} + \frac{\mathrm{e}^{-b\mu}}{\mu+1}\right).$$

If $b = +\infty$, then equality holds in both equations, implying that both $\phi_1^{(e)}(x;\lambda) = \sinh(\mu x)$ and $\phi_2^{(e)}(x;\lambda) = \cosh(\mu x)$ are eigenfunctions of the kernel for *every* $\lambda \in (0, \frac{1}{2})$.

However, if $b < +\infty$, then equality cannot hold in either equation for any value of b, so that neither $\sinh(\mu x)$ nor $\cosh(\mu x)$ are eigenfunctions of the kernel for any $\lambda \in (0, \frac{1}{2})$.

Case II: $\lambda = \frac{1}{2}$.

After substituting $\phi(x) = 1$ and then $\phi(x) = x$ into the expanded integral equation, we obtain

$$1 = 1 - \mathrm{e}^{-b} \quad \text{and} \quad x = x - (b+1)\,\mathrm{e}^{-b}\sinh x.$$

If $b = +\infty$, then inequality holds in both equations, implying that both $\phi_1^{(e)}(x;\frac{1}{2}) = 1$ and $\phi_2^{(e)}(x;\frac{1}{2}) = x$ are eigenfunctions of the kernel.

However, if $b < +\infty$, then equality cannot hold in either equation for any value of b, so that neither 1 nor x is an eigenfunction of the kernel.

Case III: $\frac{1}{2} < \lambda < +\infty$.

In this case, we set $\mu = \sqrt{2\lambda - 1}$ or $\lambda = (1+\mu^2)/2$, where $0 < \mu < +\infty$.

After substituting $\phi(x) = \sin(\mu x)$ into the expanded integral equation, we obtain

$$\sin(\mu x) = \sin(\mu x) - \mathrm{e}^{-b} \sinh x \, (\sin(\mu b) + \mu \cos(\mu b));$$

after substituting $\phi(x) = \cos(\mu x)$, we obtain

$$\cos(\mu x) = \cos(\mu x) + \mathrm{e}^{-b} \cosh x \, (\mu \sin(\mu b) - \cos(\mu b)).$$

If $b = +\infty$, then equality holds in both equations implying that both $\phi_1^{(e)}(x;\lambda) = \sin(\mu x)$ and $\phi_2^{(e)}(x;\lambda) = \cos(\mu x)$ are eigenfunctions of the kernel for *every* $\lambda \in (\frac{1}{2}, +\infty)$.

However, if $b < +\infty$, then equality holds in the first equation if and only if $\sin(\mu b) + \mu \cos(\mu b) = 0$, i.e., if $\tan(\mu b) = -\mu$. For each fixed value of b, there are an infinite number of eigenvalues $\lambda_n = (1+\mu_n^2)/2$ corresponding to the eigenfunctions $\phi_1^{(e)}(x;\lambda_n) = \sin(\mu_n x)$. Equality holds in the second equation if and only if $\mu \sin(\mu b) - \cos(\mu b) = 0$, i.e., if $\tan(\mu b) = 1/\mu$. For each fixed value of b, there are an infinite number of eigenvalues $\lambda_n = (1+\mu_n^2)/2$ corresponding to the eigenfunctions $\phi_2^{(e)}(x;\lambda_n) = \cos(\mu_n x)$.

In summary, we have come to the following conclusions:

- If $b = +\infty$, then *every* $\lambda \in (0, +\infty)$ is an eigenvalue of the kernel. Two eigenfunctions correspond to each eigenvalue λ.
- If $b < +\infty$, then there is a countably infinite set of discrete eigenvalues $\{\lambda_n\}$ which do not have a finite limit point, in perfect agreement with the Fourth Fredholm Theorem. One eigenfunction corresponds to each eigenvalue λ_n.

2.5 Constructing the Resolvent Kernel

In this section, we prove the following result:

Theorem 2.5.1. *The resolvent kernel $R(x,t;\lambda)$ can be written as the quotient of the entire functions $D(x,t;\lambda)$ and $D(\lambda)$ given by the series expansions*

$$D(x,t;\lambda) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} B_n(x,t) \lambda^n \quad \text{and} \quad D(\lambda) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} c_n \lambda^n.$$

In these series, $B_0(x,t) = K(x,t)$ and $c_0 = 1$. For every $n \geq 1$, $B_n(x,t)$ can be computed from the recursive relationship

$$B_n(x,t) = c_n K(x,t) - n \int_a^b K(x,s) B_{n-1}(s,t) \, \mathrm{d}s,$$

and c_n can be computed by evaluating the integral

$$c_n = \int_a^b B_{n-1}(t,t)\,\mathrm{d}t.$$

Before beginning the proof, we remark that the coefficients $B_n(x,t)$ and c_n in these series are computed sequentially. Schematically, the sequence of successive integrations are represented by the diagram

$$\begin{array}{ccccccc}
B_0 & \to & c_1 & & & & \\
\downarrow & \swarrow & & & & & \\
B_1 & \to & c_2 & & & & \\
\downarrow & \swarrow & & & & & \\
B_2 & \to & c_3 & & & & \\
\downarrow & \swarrow & & & & & \\
B_3 & \to & c_4 & & & & \\
\downarrow & \swarrow & & & & & \\
B_4 & \to & c_5 & & & & \\
\vdots & & \vdots & & \vdots & &
\end{array}$$

Note that if $B_N(x,t)$ vanishes identically for some N, then $c_{N+1} = 0$, and $B_n(x,t)$ also vanishes identically for all $n > N$. In this case, both series are just polynomials in λ of degree $N-1$ and N, respectively.

Proof. In order to show that the resolvent $R(x,t;\lambda)$ can be expressed as the quotient of two specific entire functions of λ, we require the following technical lemma which measures the difference between the resolvent kernels $R(x,t;\lambda)$ generated by iterating $K(x,t)$ and $R_\varepsilon(x,t;\lambda)$ generated by iterating $K_\varepsilon(x,t)$.

Lemma 2.5.1. *For all values of $\lambda \in \Delta_\rho$, we have*

$$\int_a^b U_\varepsilon(t,t;\lambda)\,\mathrm{d}t = \int_a^b R_\varepsilon(t,t;\lambda)\,\mathrm{d}t - \frac{D_\rho{}'(\lambda)}{D_\rho(\lambda)}.$$

Proof. The proof consists of a sequence of equalities requiring Eq. (2.22) in addition to the definitions of $S_\varepsilon(x,t;\lambda)$, $a_{ji}(\lambda)$, $A_{\varepsilon i}(x;\lambda)$, $a'_{ji}(\lambda)$, and the integrodifferential equation for the resolvent (see Sect. 2.3, Exercise 9). We have

$$\begin{aligned}
&\int_a^b U_\varepsilon(t,t;\lambda)\,\mathrm{d}t - \int_a^b R_\varepsilon(t,t;\lambda)\,\mathrm{d}t \\
&\quad = \int_a^b S_\varepsilon(t,t;\lambda) + \lambda \int_a^b \int_a^b S_\varepsilon(t,s;\lambda)\,R_\varepsilon(s,t;\lambda)\,\mathrm{d}s\,\mathrm{d}t \\
&\quad = \frac{1}{D_\rho(\lambda)} \sum_{i=1}^n \sum_{j=1}^n D_{ji}(\lambda) \int_a^b \left[a_i(t) + \lambda A_{\varepsilon i}(t;\lambda)\right] b_j(t)\,\mathrm{d}t
\end{aligned}$$

$$
\begin{aligned}
&+\lambda \sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)\int_a^b\int_a^b [a_i(t)+\lambda A_{\varepsilon i}(t;\lambda)]\,b_j(s)R_\varepsilon(s,t;\lambda)\,\mathrm{d}s\,\mathrm{d}t\\
&=\frac{1}{D_\rho(\lambda)}\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)\,a_{ji}(\lambda)\\
&\quad+\lambda\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)\left[\int_a^b\int_a^b R_\varepsilon(s,t;\lambda)\,a_i(t)\,b_j(s)\,\mathrm{d}t\,\mathrm{d}s\right.\\
&\quad\left.+\int_a^b\int_a^b\left(\int_a^b R_\varepsilon(s,t;\lambda)R_\varepsilon(t,u;\lambda)\,\mathrm{d}t\right)a_i(u)\,b_j(s)\,\mathrm{d}s\,\mathrm{d}u\right]\\
&=\frac{1}{D_\rho(\lambda)}\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)\,a_{ji}(\lambda)+\lambda\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)\left[\int_a^b A_{\varepsilon i}(s;\lambda)\,b_j(s)\right.\\
&\quad\left.+\lambda\int_a^b\int_a^b\frac{\partial}{\partial\lambda}R_\varepsilon(s,u;\lambda)\,a_i(u)\,b_j(s)\,\mathrm{d}u\,\mathrm{d}s\right]\\
&=\frac{1}{D_\rho(\lambda)}\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)\,a_{ji}(\lambda)\\
&\quad+\lambda\,\frac{1}{D_\rho(\lambda)}\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)\\
&\quad\times\int_a^b\left(A_{\varepsilon i}(s;\lambda)+\lambda\,\frac{\partial}{\partial\lambda}\left(\int_a^b R_\varepsilon(s,u;\lambda)\,a_i(u)\,\mathrm{d}u\right)\right)b_j(s)\,\mathrm{d}s\\
&=\frac{1}{D_\rho(\lambda)}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)\left(a_{ji}(\lambda)+\lambda a'_{ji}(\lambda)\right)\right]\\
&=-\frac{1}{D_\rho(\lambda)}\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)\,(-\lambda a_{ji}(\lambda))'\\
&=-\frac{D_\rho{}'(\lambda)}{D_\rho(\lambda)}.
\end{aligned}
$$

□

This lemma will be used to show that there exist entire functions $D(x,t;\lambda)$ and $D(\lambda)$ such that

$$D(x,t;\lambda)=R(x,t;\lambda)\,D(\lambda).$$

Since $R(t,t;\lambda)$ is an analytic function of λ in Δ_ρ, the statement of the lemma implies that the zeroes of $D_\rho(\lambda)$ correspond to the poles of the function

$$\delta_\varepsilon(\lambda)=\int_a^b U_\varepsilon(t,t;\lambda)\,\mathrm{d}t$$

that lie in Δ_ρ. Specifically, if $\lambda_k \in \Delta_\rho$ is a zero of algebraic multiplicity m_k in $D_\rho(\lambda)$, then $D_\rho(\lambda) = (\lambda - \lambda_k)^{m_k} e_\rho(\lambda)$, where $e_\rho(\lambda) \neq 0$ near λ_k, and

$$\frac{D_\rho{}'(\lambda)}{D_\rho(\lambda)} = \frac{m_k}{\lambda - \lambda_k} + \frac{e_\rho{}'(\lambda)}{e_\rho(\lambda)}.$$

This fact in conjunction with the equality stated in the lemma shows that λ_k is a simple pole of $\delta_\varepsilon(\lambda)$ with residue $-m_k$. Hence, $\delta_\varepsilon(\lambda)$ has only simple poles which coincide with the zeroes of $D_\rho(\lambda)$. Since $U_\varepsilon(t,t;\lambda)$ is the restriction of $R(t,t;\lambda)$ to the disk Δ_ρ, $\delta_\varepsilon(\lambda)$ is also the restriction of

$$\delta(\lambda) = \int_a^b R(t,t;\lambda)\,\mathrm{d}t$$

to the disk Δ_ρ, so that $\delta(\lambda)$ is a meromorphic function in the full λ-plane having only simple poles that coincide with the eigenvalues of the kernel $K(x,t)$. From this, it follows that

$$D(\lambda) = \exp\left(-\int_0^\lambda \delta(\lambda)\,\mathrm{d}\lambda\right)$$

is an entire function of λ.

To show that $D(x,t;\lambda)$ is entire, note that we can integrate the equality stated in the lemma near the origin since $D_\rho(0) = 1$ to obtain the relation

$$-\int_0^\lambda \delta(\lambda)\,\mathrm{d}\lambda = -\int_0^\lambda \int_a^b R_\varepsilon(t,t;\lambda)\,\mathrm{d}t\,\mathrm{d}\lambda + \ln D_\rho(\lambda)$$

which is valid in Δ_ρ. Upon exponentiation, we obtain

$$D(\lambda) = D_\rho(\lambda) \cdot \exp\left(-\int_0^\lambda \int_a^b R_\varepsilon(t,t;\lambda)\,\mathrm{d}t\,\mathrm{d}\lambda\right).$$

It follows from this relation that the zeroes of $D(\lambda)$ coincide with those of $D_\rho(\lambda)$ in Δ_ρ and have the same multiplicity, since the exponential factor does not vanish. Thus, $D(x,t;\lambda)$ is an entire function of λ, since the poles of $R(t,t;\lambda)$ are cancelled by the zeroes of $D(\lambda)$.

It remains to develop the series expansions in the statement of the theorem. Recall from Sect. 2.3 that the resolvent kernel can be represented by the expansion

$$R(x,t;\lambda) = \sum_{m=1}^{\infty} \lambda^{m-1} K_m(x,t)$$

for small values of λ, where $K_m(x,t)$ is the m-th iterated kernel of $K(x,t)$. It follows that

$$\delta(\lambda) = \int_a^b R(t,t;\lambda)\,\mathrm{d}t = \sum_{m=1}^{\infty} A_m \lambda^{m-1},$$

where we have set

$$A_m = \int_a^b K_m(t,t)\,\mathrm{d}t.$$

The coefficient A_m is known as the *trace of the kernel* $K_m(x,t)$. It follows that

$$\begin{aligned}
D(\lambda) &= \exp\left[-\int_0^\lambda \left(\sum_{m=1}^{\infty} A_m \lambda^{m-1}\right)\mathrm{d}\lambda\right] \\
&= \exp\left(-\sum_{m=1}^{\infty}\frac{A_m}{m}\lambda^m\right) \\
&= \sum_{k=0}^{\infty}\frac{(-1)^k}{k!}\left(\sum_{m=1}^{\infty}\frac{A_m}{m}\lambda^m\right) \\
&= 1+\frac{(-1)^1}{1!}A_1\lambda+\frac{(-1)^2}{2!}\left({A_1}^2-A_2\right)\lambda^2 \\
&\quad+\frac{(-1)^3}{3!}\left({A_1}^3-3A_1A_2+2A_3\right)\lambda^3 \\
&\quad+\frac{(-1)^4}{4!}\left({A_1}^4-6{A_1}^2A_2+3{A_2}^2+8A_1A_3-6A_4\right)\lambda^4+\cdots.
\end{aligned}$$

More simply, we can write

$$D(\lambda) = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}c_n\lambda^n,$$

where $c_0 = 1$. Also, for $n \geq 2$, each coefficient c_n is a evidently a multivariable polynomial in the traces $A_1,\ldots,A_n$ whose coefficients are integers that sum to zero. (To establish this purely algebraic statement, just set $A_m = 1$ for $m \geq 1$. Then $D(\lambda) = 1-\lambda$, so that $c_n = 0$ for all $n \geq 2$.)

Since $D(x,t;\lambda)$ is entire, it has the Maclaurin expansion

$$D(x,t;\lambda) = \sum_{n=0}^{\infty}\frac{1}{n!}D^{(n)}(x,t;0)\lambda^n = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}B_n(x,t)\lambda^n.$$

The coefficients $B_n(x,t)$ in this series must be computed in order to establish the recursion relation for them given in the statement of the theorem. If $n = 0$, then

$$B_0(x,t) = D(x,t;0) = R(x,t;0)D(0) = K(x,t).$$

If $n \geq 1$, then we apply Leibniz formula to obtain

$$\begin{aligned}
B_n(x,t) &= (-1)^n D^{(n)}(x,t;0) \\
&= (-1)^n\left(D(0)R(x,t;0)\right)^{(n)}
\end{aligned}$$

$$
\begin{aligned}
&= (-1)^n \sum_{m=0}^{n} \frac{n!}{m!\,(n-m)!}\, D^{(n-m)}(0)\, R^{(m)}(x,t;0) \\
&= \sum_{m=0}^{n} \frac{n!}{m!\,(n-m)!}\, \left((-1)^{-m} c_{n-m}\right) \left(m!\, K_{m+1}(x,t)\right) \\
&= n! \sum_{m=0}^{n} (-1)^{-m} \frac{1}{(n-m)!}\, c_{n-m}\, K_{m+1}(x,t) \\
&= n! \left(\frac{1}{n!}\, c_n K(x,t) + \sum_{m=1}^{n} \frac{(-1)^{-m}}{(n-m)!}\, c_{n-m}\, K_{m+1}(x,t) \right) \\
&= c_n K(x,t) - n \left((n-1)! \sum_{m=0}^{n-1} \frac{(-1)^{-m}}{((n-1)-m)!}\, c_{(n-1)-m}\, K_{m+2}(x,t) \right) \\
&= c_n K(x,t) \\
&\quad - n \left[\int_a^b K(x,s) \right. \\
&\quad \left. \times \left((n-1)! \sum_{m=0}^{n-1} \frac{(-1)^{-m}}{((n-1)-m)!}\, c_{(n-1)-m}\, K_{m+1}(s,t) \right) \mathrm{d}s \right] \\
&= c_n K(x,t) - n \int_a^b K(x,s)\, B_{n-1}(s,t)\, \mathrm{d}s.
\end{aligned}
$$

This sequence of equalities shows that $B_n(x,t)$ is a particular linear combination of iterated kernels and that $B_n(x,t)$ can be computed recursively.

Finally, we show that each c_{n+1} may be determined from the coefficient function $B_n(x,t)$. On the one hand, we have

$$
\begin{aligned}
\int_a^b D(t,t;\lambda)\, \mathrm{d}t &= D(\lambda) \int_a^b R(t,t;\lambda)\, \mathrm{d}t \\
&= D(\lambda)\, \delta(\lambda) \\
&= -D'(\lambda) \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!}\, c_{n+1}\, \lambda^n.
\end{aligned}
$$

On the other hand, we also have

$$
\int_a^b D(t,t;\lambda)\, \mathrm{d}t = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \left(\int_a^b B_n(t,t)\, \mathrm{d}t \right) \lambda^n.
$$

By comparing the coefficients of the powers of λ in these two expansions, we have

$$c_{n+1} = \int_a^b B_n(t,t)\,\mathrm{d}t$$

for all $n \geq 0$ as required. □

The coefficients c_n are related to the traces A_m in the following way. By setting $x = t$ in the sequence of equalities in the proof, we obtain the relation

$$B_n(t,t) = n!\left(\sum_{m=0}^{n}(-1)^m \frac{1}{(n-m)!}\,c_{n-m}K_{m+1}(t,t)\right).$$

Upon integration, we immediately have

$$c_{n+1} = n!\left(\sum_{m=0}^{n}(-1)^m \frac{1}{(n-m)!}\,c_{n-m}A_{m+1}\right).$$

It is of interest to note that the coefficients c_n can be written as integrals of determinants. By definition, we have

$$c_1 = A_1 = \int_a^b K(t,t)\,\mathrm{d}t.$$

Similarly, we have

$$\begin{aligned}
c_2 &= {A_1}^2 - A_2 \\
&= \left(\int_a^b K(t_1,t_1)\,\mathrm{d}t_1\right)^2 - \int_a^b K_2(t_1,t_1)\,\mathrm{d}t_1 \\
&= \left(\int_a^b K(t_1,t_1)\,\mathrm{d}t_1\right)\left(\int_a^b K(t_2,t_2)\,\mathrm{d}t_2\right) - \int_a^b\int_a^b K(t_1,t_2)K(t_2,t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1 \\
&= \int_a^b\int_a^b \det\begin{pmatrix} K(t_1,t_1) & K(t_1,t_2) \\ K(t_2,t_1) & K(t_2,t_2)\end{pmatrix}\mathrm{d}t_2\,\mathrm{d}t_1.
\end{aligned}$$

In general, it can be shown that

$$c_n = \int_a^b \cdots \int_a^b \det\begin{pmatrix} K(t_1,t_1) & K(t_1,t_2) & \cdots & K(t_1,t_n) \\ K(t_2,t_1) & K(t_2,t_2) & \cdots & K(t_1,t_n) \\ \vdots & \vdots & \ddots & \vdots \\ K(t_n,t_1) & K(t_n,t_2) & \cdots & K(t_n,t_n)\end{pmatrix}\mathrm{d}t_n \ldots \mathrm{d}t_1.$$

The magnitude of the coefficients $|c_n|$ can be estimated with a classical inequality due to Hadamard stating that the absolute value of the determinant of a matrix is less than or equal to the product of the norms of its rows. More precisely, if $\mathbf{A} = (a_{ij})$ is an $n \times n$ matrix and $\mathbf{a}_i = (a_{i1}, \ldots, a_{in})$ is its ith row, then

$$|\det \mathbf{A}| \leq \prod_{i=1}^{n} \|\mathbf{a}_i\|,$$

where $\|\mathbf{a}_i\| = (a_{i1}^2 + \cdots + a_{in}^2)^{1/2}$. By Hadamard's inequality, we have

$$|c_n| \leq n^{n/2} M^n (b-a)^n,$$

if $|K(x,t)| \leq M$ on the square $Q(a,b)$.

The coefficient functions $B_n(x,t)$ can also be written as multiple integrals of a determinant. Specifically, for $n \geq 1$, we have

$$B_n(x,t) = \int_a^b \ldots \int_a^b \Delta_n(x,t)\, \mathrm{d}t_n, \ldots, \mathrm{d}t_1,$$

where

$$\Delta_n(x,t) = \det \begin{pmatrix} K(x,t) & K(x,t_1) & K(x,t_2) & \ldots & K(x,t_n) \\ K(t_1,t) & K(t_1,t_1) & K(t_1,t_2) & \ldots & K(t_1,t_n) \\ K(t_2,t) & K(t_2,t_1) & K(t_2,t_2) & \ldots & K(t_1,t_n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ K(t_n,t) & K(t_n,t_1) & K(t_n,t_2) & \ldots & K(t_n,t_n) \end{pmatrix}.$$

Illustrative Examples

- *Example 1*: Recall the inhomogeneous Fredholm integral equation of the second kind

$$\phi(x) = f(x) + \lambda \int_0^1 (xt^2 + x^2 t^4)\, \phi(t)\, \mathrm{d}t$$

that was solved in Sect. 1.2. It is always the case that $c_0 = 1$. Since $B_0(x,t) = K(x,t) = xt^2 + x^2 t^4$, we have

$$c_1 = \int_0^1 B_0(t,t)\, \mathrm{d}t = \int_0^1 (t^3 + t^6)\, \mathrm{d}t = \tfrac{11}{28}.$$

Also,

$$\begin{aligned} B_1(x,t) &= c_1 K(x,t) - 1 \cdot \int_0^1 K(x,s) B_0(s,t)\,\mathrm{d}t \\ &= \tfrac{11}{28}(xt^2 + x^2t^4) - \int_0^1 (xs^2 + x^2s^4)(st^2 + s^2t^4)\,\mathrm{d}t \\ &= \tfrac{1}{4}x^2t^4 + \tfrac{1}{7}xt^2 - \tfrac{1}{5}xt^4 - \tfrac{1}{6}x^2t^2 \end{aligned}$$

and

$$c_2 = \int_0^1 B_1(t,t)\,\mathrm{d}t = \int_0^1 (\tfrac{1}{4}t^6 + \tfrac{1}{7}t^3 - \tfrac{1}{5}t^5 - \tfrac{1}{6}t^4)\,\mathrm{d}t = \tfrac{1}{210}.$$

One more integration shows that $B_2(x,t) \equiv 0$, so that $c_n = 0$ and $B_n(x,t) \equiv 0$ if $n \geq 3$. Hence,

$$D(\lambda) = 1 + (-1)\frac{1}{1!}\left(\tfrac{11}{28}\right)\lambda + (-1)^2\frac{1}{2!}\left(\tfrac{1}{210}\right)\lambda^2 = \tfrac{1}{420}\left(\lambda^2 - 165\lambda + 420\right)$$

and

$$\begin{aligned} D(x,t;\lambda) &= B_0(x,t) + (-1)\frac{1}{1!}B_1(x,t)\lambda \\ &= (xt^2 + x^2t^4) - \left(\tfrac{1}{7}xt^2 + \tfrac{1}{4}x^2t^4 - \tfrac{1}{5}xt^4 - \tfrac{1}{6}x^2t^2\right)\lambda. \end{aligned}$$

These results are identical to those obtained in Sect. 1.2.

Section 2.5 Exercises

1. Let $K(x,t) = a(x)\overline{b(t)}$, and assume that $\alpha = \int_a^b a(t)\overline{b(t)}\,\mathrm{d}t \neq 0$. Show that

$$R(x,t;\lambda) = \frac{K(x,t)}{1 - \lambda\alpha}$$

and that the mth trace $A_m = \alpha^m$. In particular, if $b(t) = a(t)$, then $A_m = \|a\|_2^{2m}$.

2. Let $K(x,t) = a_1(x)\overline{b_1(t)} + a_2(x)\overline{b_2(t)}$ and define the matrix $\mathbf{A} = (a_{ij})$ by

$$a_{ij} = \int_a^b a_j(t)\overline{b_i(t)}\,\mathrm{d}t.$$

Determine the resolvent kernel $R(x,t;\lambda)$ and show that

$$\int_a^b R(t,t;\lambda)\,\mathrm{d}t = \frac{\mathrm{tr}(\mathbf{A}) - 2\det(\mathbf{A})\lambda}{1 - \mathrm{tr}(\mathbf{A})\lambda + \det(\mathbf{A})\lambda^2} = \sum_{m=1}^{\infty} A_m \lambda^{m-1}$$

if $|\lambda| < \min\{|\lambda_1|, |\lambda_2|\}$. Thus, the mth trace A_m can be expressed in terms of $\mathrm{tr}(\mathbf{A})$ and $\det(\mathbf{A})$.

Also, since

$$D(\lambda) = 1 - \mathrm{tr}(\mathbf{A})\,\lambda + \det(\mathbf{A})\,\lambda^2 = \left(1 - \frac{\lambda}{\lambda_1}\right)\left(1 - \frac{\lambda}{\lambda_2}\right),$$

we have

$$\mathrm{tr}(\mathbf{A}) = \frac{1}{\lambda_1} + \frac{1}{\lambda_2} \quad \text{and} \quad \det(\mathbf{A}) = \frac{1}{\lambda_1\,\lambda_2}.$$

Use this observation to show that

$$A_m = \frac{1}{\lambda_1^m} + \frac{1}{\lambda_2^m}$$

for every $m \geq 1$.

2.6 Numerical Methods

We have seen that considerable difficulties may arise in the computation of the resolvent kernel, even with relatively simple continuous general kernels. Thus, the use of numerical methods assumes critical importance in the production of approximate solutions to Fredholm integral equations of the second kind. In this section, we consider several elementary approaches to the problem of determining approximations to the solution of an integral equation.

2.6.1 The Method of Uniform Approximation

The essence of the theorem in this subsection is that if the two free terms and the two kernels in two Fredholm integral equations are close, then their solutions are close as well. This theorem is extremely useful if one of the integral equations is difficult to solve while the other one is not.

Theorem 2.6.1 (Theorem of Uniform Approximation). *Let $f(x)$ and $g(x)$ be continuous free terms defined on the interval $[a,b]$, and let $K(x,t)$ and $L(x,t)$ be continuous kernels defined on the square $Q(a,b)$.*

Suppose that

$$\max_{a\leq x\leq b} |f(x) - g(x)| < \varepsilon \quad \textit{and} \quad \max_{a\leq x,t\leq b} |K(x,t) - L(x,t)| < \kappa$$

for some fixed constants ε and κ. Suppose further that λ is a common regular value of both kernels of the Fredholm integral equations

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t$$

and

$$\psi(x) = g(x) + \lambda \int_a^b L(x,t)\,\psi(t)\,\mathrm{d}t$$

for which $|\lambda|$ *is less than the radii of convergence of both of the resolvent series* $R_K(x,t;\lambda)$ *and* $R_L(x,t;\lambda)$.

Then there exist constants $\alpha = \alpha(\lambda, B_L)$ *and* $\beta = \beta(\lambda, B_f, B_K, B_L)$ *such that*

$$\max_{a\le x\le b} |\phi(x) - \psi(x)| \le \alpha\,\varepsilon + \beta\,\kappa,$$

where

$$B_f = \max_{a\le x\le b} |f(x)|,$$

$$B_K(\lambda) = \max_{a\le x\le b} \left| \int_a^b R_K(x,t;\lambda)\,\mathrm{d}t \right|,$$

and

$$B_L(\lambda) = \max_{a\le x\le b} \left| \int_a^b R_L(x,t;\lambda)\,\mathrm{d}t \right|.$$

Proof. By virtue of the First Fredholm Theorem in Sect. 2.4, the solution to the first equation in the statement of the theorem has the representation

$$\phi(x) = f(x) + \lambda \int_a^b R_K(x,t;\lambda)\,f(t)\,\mathrm{d}t.$$

It follows from this representation that

$$B_\phi = \max_{a\le x\le b} |\phi(x)| \le B_f\,(1 + |\lambda|\,B_K(\lambda)).$$

Next, if we subtract the two integral equations given in the statement of the theorem, we obtain

$$\phi(x) - \psi(x) = D(x) + \lambda \int_a^b L(x,t)\,(\phi(t) - \psi(t))\,\mathrm{d}t,$$

where

$$D(x) = f(x) - g(x) + \lambda \int_a^b (K(x,t) - L(x,t))\,\phi(t)\,\mathrm{d}t.$$

It will be useful to have the estimate

$$B_D = \max_{a\le x\le b} |D(x)| \le \varepsilon + |\lambda|\,\kappa\,(b-a)\,B_\phi.$$

Again, by the First Fredholm Theorem, the solution to the above equation has the representation

$$\phi(x) - \psi(x) = D(x) + \lambda \int_a^b R_L(x,t;\lambda)\, D(t)\, \mathrm{d}t.$$

It follows from this representation that

$$\max_{a \le x \le b} |\phi(x) - \psi(x)| \le (1 + |\lambda|\, B_L(\lambda))\, B_D.$$

Upon combining the above estimates, we finally obtain

$$\max_{a \le x \le b} |\phi(x) - \psi(x)| \le \alpha\, \varepsilon + \beta\, \kappa$$

where

$$\alpha = 1 + |\lambda|\, B_L(\lambda)$$

and

$$\beta = |\lambda|\, (1 + |\lambda|\, B_L(\lambda))\, (1 + |\lambda|\, B_K(\lambda))\, (b-a) B_f. \qquad \square$$

In order to use this theorem, it is necessary to obtain estimates for $B_K(\lambda)$ and $B_L(\lambda)$. If $K(x,t) \le M$ on the square $Q(a,b)$ and λ is within the radius of convergence of the resolvent series, then

$$\left| \int_a^b R_K(x,t;\lambda)\, \mathrm{d}t \right| \le \frac{M(b-a)}{1 - |\lambda| M(b-a)} = B_K(\lambda).$$

Illustrative Examples

- *Example 1*: Consider the integral equation

$$\phi(x) = \cos x + \tfrac{1}{2} \int_0^1 \cos(xt)\, \phi(t)\, \mathrm{d}t.$$

The kernel $K(x,t) = \cos(xt)$ is not separable, and it is difficult to compute its iterated kernels. Consequently, an approximative method is indicated.

Since the free term $\cos x$ can be represented with the Taylor series

$$\cos x = 1 - \frac{1}{2!} x^2 + \frac{1}{4!} x^4 - \frac{1}{6!} x^6 + \cdots,$$

we choose

$$g(x) = 1 - \frac{1}{2!} x^2 + \frac{1}{4!} x^4 \quad \text{and} \quad L(x,t) = 1 - \frac{1}{2!} x^2 t^2 + \frac{1}{4!} x^4 t^4$$

and consider the companion equation

$$\psi(x) = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \frac{1}{2}\int_0^1 \left(1 - \frac{1}{2!}x^2t^2 + \frac{1}{4!}x^4t^4\right)\psi(t)\,\mathrm{d}t.$$

The companion equation has a separable kernel. It can be solved by employing the method prescribed in Sect. 1.3 to obtain

$$\psi(x) = \frac{11,532,090}{6,397,711} - \frac{7,944,195}{12,795,422}x^2 + \frac{607,005}{12,795,422}x^4.$$

If we substitute this approximation into the equation, then we obtain

$$\begin{aligned}
&\psi(x) - \cos x - \frac{1}{2}\int_0^1 \cos(xt)\,\psi(t)\,\mathrm{d}t \\
&\quad = \frac{1,620,362,251}{1,064,067,293,520}x^6 - \frac{1,473,509,027}{55,331,499,263,040}x^8 + O(x^9) \\
&\quad \approx 0.0015228x^6 - 0.0000266x^8
\end{aligned}$$

as the residual. The constant term and the coefficients of x^2 and x^4 in this difference vanish.

The actual error is well within the error predicted by the theorem. Since the Taylor series for $\cos x$ is alternating, we have the classical estimates

$$|f(x) - g(x)| = \left|\cos x - \left(1 - \frac{1}{2}x^2 + \frac{1}{24}x^4\right)\right| < \varepsilon = \frac{1}{720}$$

on the interval $[0,1]$ and

$$|K(x,t) - L(x,t)| = \left|\cos(xt) - \left(1 - \frac{1}{2}x^2t^2 + \frac{1}{24}x^4t^4\right)\right| < \kappa = \frac{1}{720}$$

on the square $Q(0,1)$. Also, $B_f = 1$ and $B_K(\frac{1}{2}) = B_L(\frac{1}{2}) = 2$. Thus, the difference between the actual and the approximate solutions is predicted by the theorem to be no more than $|\phi(x) - \psi(x)| \leq \frac{1}{180} = 0.0055555$.

2.6.2 The Method of Collocation

The Method of Collocation produces a function $y^*(x)$ that approximates the solution $\phi(x)$ to a Fredholm integral equation on an interval as a combination of a predetermined set of continuous and linearly independent functions on that interval.

Although this method is described here for Fredholm integral equations, we note that it can be adapted to produce an approximation to the solution of other types of integral equations as well.

Theorem 2.6.2 (The Collocation Theorem). *Consider the integral equation*

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t,$$

where the free term $f(x)$ is continuous on the interval $[a,b]$, the kernel $K(x,t)$ is continuous on the square $Q(a,b)$, and λ is a regular value of the kernel that is less than the radius of convergence of the resolvent series.

Choose a selection $\{x_1,\ldots,x_n\}$ of n nodes with $a \le x_1 < \cdots < x_n \le b$ and a set $\{y_1(x),\ldots,y_n(x)\}$ of n functions that are continuous and linearly independent on the interval $[a,b]$. Define the matrices $\mathbf{Y} = (y_{ik})$ and $\mathbf{Z} = (z_{ik})$ where

$$y_{ik} = y_k(x_i) \quad \textit{and} \quad z_{ik} = \int_a^b K(x_i,t)\,y_k(t)\,\mathrm{d}t.$$

If $\det(\mathbf{Y} - \lambda\,\mathbf{Z}) \neq 0$*, then there exists a linear combination*

$$y^*(x) = \sum_{i=1}^{n} a_i^*\,y_i(x)$$

for which

$$y^*(x_i) = f(x_i) + \lambda \int_a^b K(x_i,t)\,y^*(t)\,\mathrm{d}t$$

for all $i = 1,\ldots,n$. Furthermore, the inequality

$$|\phi(x) - y^*(x)| \le (1 + |\lambda|\,B_R(\lambda))\,B_{y^*}$$

holds uniformly on the interval $[a,b]$, where

$$B_R(\lambda) = \max_{a\le x\le b} \left| \int_a^b R_K(x,t;\lambda)\,\mathrm{d}t \right|$$

and

$$B_{y^*} = \max_{a\le x\le b} \left| y^*(x) - f(x) - \lambda \int_a^b K(x,t)\,y^*(t)\,\mathrm{d}t \right|.$$

Proof. For any choice of $y(x)$, define the *residual* $\eta(x)$ *of* $y(x)$ to be

$$\eta(x) = y(x) - f(x) - \lambda \int_a^b K(x,t)\,y(t)\,\mathrm{d}t.$$

The residual $\eta(x) \equiv 0$ on $[a,b]$ if and only if $y(x)$ is equal to the unique solution $\phi(x)$ of the integral equation.

Let $\{a_1,\ldots,a_n\}$ be an arbitrary choice of coefficents and consider the linear combination

$$y(x) = y(x;a_1,\ldots,a_n) = \sum_{k=1}^{n} a_k y_k(x).$$

For this choice of coefficients, the residual of $y(x)$ assumes the corresponding form

$$\eta(x) = \eta(x;a_1,\ldots,a_n) = \sum_{k=1}^{n} a_k \left(y_k(x) - \lambda z_k(x)\right) - f(x).$$

If $y(x)$ is not the solution to the integral equation, then the residual $\eta(x)$ does not identically vanish on the interval $[a,b]$. However, if $\eta(x_i) = 0$ for all $i = 1,\ldots,n$, then the approximate solution $y(x)$ will be measureably close to the actual solution $\phi(x)$ on the entire interval $[a,b]$. It is this idea which constitutes the essence of the Method of Collocation.

By requiring that $\eta(x_i) = 0$ for all $i = 1,\ldots,n$, we produce a system of n linear equations in the n unknowns $a_1,\ldots,a_n$. This linear system may be written in matrix form as

$$(\mathbf{Y} - \lambda\,\mathbf{Z})\,\mathbf{a} = \mathbf{f}$$

where $\mathbf{a} = (a_1,\ldots,a_n)^{\mathrm{T}}$ and $\mathbf{f} = (f(x_1),\ldots,f(x_n))^{\mathrm{T}}$. If $\det(\mathbf{Y} - \lambda\,\mathbf{Z}) \neq 0$, then the linear system has a unique solution

$$\mathbf{a}^* = (a_1^*,\ldots,a_n^*)^{\mathrm{T}} = (\mathbf{Y} - \lambda\,\mathbf{Z})^{-1}\,\mathbf{f}.$$

The corresponding approximate solution

$$y^*(x) = y^*(x;a_1^*,\ldots,a_n^*) = \sum_{k=1}^{n} a_k^* y_k(x)$$

has the property that its residual vanishes at the selected nodes x_i, i.e., that $\eta^*(x_i) = \eta^*(x_i;a_1^*,\ldots,a_n^*) = 0$ for all $i = 1,\ldots,n$.

For any choice of $y(x)$, let $\varepsilon(x) = y(x) - \phi(x)$. By adding the integral equation to the definition of the residual of $y(x)$, we obtain

$$\varepsilon(x) = \eta(x) + \lambda \int_a^b K(x,t)\,\varepsilon(t)\,\mathrm{d}t.$$

Since $\eta(x)$ continuous on the interval $[a,b]$, this equation has a unique solution that can be represented in the form

$$\varepsilon(x) = \eta(x) + \lambda \int_a^b R(x,t;\lambda)\,\eta(t)\,\mathrm{d}t,$$

where $R(x,t;\lambda)$ is the unique resolvent kernel corresponding to the given kernel $K(x,t)$.

The estimate

$$|\varepsilon(x)| = |y(x) - \phi(x)| \leq (1 + |\lambda| B_R(\lambda)) B_y$$

follows directly from the integral representation for $\varepsilon(x)$, where

$$B_y = \max_{a \leq x \leq b} |\eta(x)|$$

and $B_R(\lambda)$ is a bound on the integral of the resolvent. Thus, the magnitude of the residual controls the accuracy of the estimation. Since this estimate holds for the approximant $y^*(x)$, the proof is complete. □

Some comments regarding the Method of Collocation may be helpful.

To employ this method, it must be possible to compute the integrals $z_k(x)$, which depend upon the choice of the functions $y_k(x)$. Some typical choices for the functions $y_k(x)$ include various types of polynomials, such as a finite set of powers of x (as illustrated in the example below), Legendre or Chebychev polynomials, or trigonometric polynomials. The best choice depends upon the functional form of the kernel $K(x,t)$.

If the determinant $\det(\mathbf{Y} - \lambda \mathbf{Z}) = 0$, then an alternate choice of nodes and/or basis functions is required. This situation might arise if the basis functions were accidentally chosen to be eigenfunctions of the kernel. If the determinant does not vanish, but is close to zero, then computational difficulties may arise, since the entries in $\mathbf{Y} - \lambda \mathbf{Z}$ might be rather large.

Requiring that $\eta(x_i) = 0$ for all $i = 1, \ldots, n$ is not the only method of producing a small residual. For example, we could require that the coefficients $a_1, \ldots, a_n$ are such that the sum

$$\sum_{i=1}^{n} |\eta(x_i; a_1, \ldots, a_n)|^2$$

is minimal. In this case, the bound on the residual would still control the accuracy of the approximation.

An examination of the graph of the residual $|\eta(x)|$ can lead to improved accuracy for the approximate solution. For if the graph of the residual has a few maxima, but is otherwise small, it is reasonable to suggest that shifting the nodes, i.e., *adapting the mesh*, could lead to a decrease in the magnitude of the residual.

The bound $B_R(\lambda)$ can be chosen to be

$$B_R(\lambda) = \frac{B_K (b-a)}{1 - |\lambda| B_K (b-a)},$$

provided that $|\lambda|$ is less than the radius of convergence of the resolvent series. For other values of $|\lambda|$, an alternate estimate must be made.

A Concise Guide to Computation

- *Preliminary calculations*:
  1. Choose a set $\{x_1,\ldots,x_n\}$ of n nodes with $a \le x_1 < \cdots < x_n \le b$. Usually, the nodes are chosen to be equally spaced, i.e.,
  $$\Delta x_i = x_i - x_{i-1} = \frac{b-a}{n}$$
  for all $i = 1,\ldots,n$.
  2. Choose a set $\{y_1(x),\ldots,y_n(x)\}$ of n functions that are continuous and linearly independent on the interval $[a,b]$.
  3. Compute the elements of the matrices $\mathbf{Y} = (y_{ik})$ and $\mathbf{Z} = (z_{ik})$, where
  $$y_{ik} = y_k(x_i) \quad \text{and} \quad z_{ik} = \int_a^b K(x_i,t)\,y_k(t)\,\mathrm{d}t$$
  for all $i,k = 1,\ldots,n$.
  4. Compute the vector $\mathbf{f} = (f(x_i))$.
- *Determine the approximate solution*:
  1. Compute the determinant $\det(\mathbf{Y} - \lambda\,\mathbf{Z})$. If this determinant vanishes, then adjust the choices of nodes and/or basis functions as necessary.
  2. If the matrix $\mathbf{Y} - \lambda\,\mathbf{Z}$ is invertible, then solve the matrix equation
  $$(\mathbf{Y} - \lambda\,\mathbf{Z})\,\mathbf{a} = \mathbf{f}$$
  for its unique solution
  $$\mathbf{a}^* = (a_1^*,\ldots,a_n^*)^{\mathrm{T}} = (\mathbf{Y} - \lambda\,\mathbf{Z})^{-1}\,\mathbf{f}.$$
  3. Use the solution to formulate the approximation
  $$y^*(x) = \sum_{i=1}^{n} a_i^*\,y_i(x).$$
- *Estimate the accuracy of the approximate solution*:
  1. Formulate the residual $\eta^*(x)$ in terms of $y^*(x)$.
  2. Determine the upper bounds B_{y^*} and $B_K(\lambda)$.
  3. Use these computed bounds to compute an estimate for the upper bound on the possible error $|\varepsilon(x)|$.
  4. If the error must be reduced, then plot the residual to determine an appropriate number of additional nodes and/or basis functions, and repeat the previous steps as appropriate.

Illustrative Examples

- *Example 1*: Consider the Fredholm integral equation

$$\phi(x) = \sin(\pi x) + \lambda \int_0^1 \frac{1}{1+x+t}\,\phi(t)\,\mathrm{d}t.$$

According to the Theorem of Successive Approximation, there are no eigenvalues of the kernel in the disk of radius $1/\|K\|_2 = 1.864419\ldots$, where

$$\|K\|_2 = \left(\int_0^1 \int_0^1 \left(\frac{1}{1+x+t}\right)^2 \mathrm{d}t\,\mathrm{d}x\right)^{1/2} = \sqrt{\ln(\tfrac{4}{3})} = 0.536360\ldots.$$

Therefore, the choice $\lambda = \frac{1}{2}$ is a regular value of the kernel.

Suppose that we desire a polynomial approximation of the sixth degree to the solution $\phi(x)$ of the equation. Following the guide, we choose the seven equally spaced nodes $\{0, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{5}{6}, 1\}$ and the seven basis functions $\{1, x, x^2, x^3, x^4, x^5, x^6\}$. The approximate solution has the form

$$y(x) = \sum_{k=0}^{6} a_k\, y_k(x) = \sum_{k=0}^{6} a_k x^k,$$

and the residual $\eta(x)$ assumes the form

$$\eta(x) = \sum_{k=0}^{6} a_k\,(x^k - \tfrac{1}{2} z_k(x)) - \sin(\pi x).$$

Requiring that $\eta(x_i) = 0$ for $i = 0, \ldots, 6$ leads to the 7×7 matrix system

$$(\mathbf{Y} - \tfrac{1}{2}\,\mathbf{Z})\,\mathbf{a} = \mathbf{f},$$

where $(\mathbf{Y} - \frac{1}{2}\,\mathbf{Z})_{ik} = (x_i^k - \frac{1}{2}\,z_k(x_i))$ and $\mathbf{f} = (0, \frac{1}{2}, \frac{\sqrt{3}}{2}, 1, \frac{\sqrt{3}}{2}, \frac{1}{2}, 0)^{\mathrm{T}}$. With rounding in the last decimal place, the matrix $\mathbf{Y} - \frac{1}{2}\,\mathbf{Z}$ is given by

$$\begin{pmatrix}
+0.653426 & -0.153426 & -0.096574 & -0.070093 & -0.054907 & -0.045093 & -0.038240 \\
+0.690480 & +0.027773 & -0.060180 & -0.059420 & -0.049504 & -0.041217 & -0.035075 \\
+0.720192 & +0.206411 & +0.030342 & -0.021937 & -0.034023 & -0.034061 & -0.031061 \\
+0.744587 & +0.383119 & +0.175321 & +0.070352 & +0.019473 & -0.004209 & -0.014520 \\
+0.764998 & +0.558336 & +0.374995 & +0.245379 & +0.157393 & +0.098583 & +0.059632 \\
+0.782341 & +0.732375 & +0.629535 & +0.531038 & +0.444641 & +0.370834 & +0.308478 \\
+0.797267 & +0.905465 & +0.939070 & +0.955194 & +0.964613 & +0.970775 & +0.975116
\end{pmatrix}.$$

After determining the solution $\mathbf{a}^* = (\mathbf{Y} - \frac{1}{2}\mathbf{Z})^{-1}\mathbf{f}$ to this system, we can formulate the approximate solution

$$\begin{aligned} y^*(x) &= y^*(x; a_1^*, \ldots, a_n^*) \\ &= 0.297340 + 2.932409\,x + 0.124667\,x^2 - 5.092254\,x^3 \\ &\quad - 0.557749\,x^4 + 3.710938\,x^5 - 1.240580\,x^6. \end{aligned}$$

Inspecting the graph of the residual $\eta^*(x)$ shows that $\eta^*(x_i) = 0$ and that $|\eta^*(x)| \leq 0.000036$. With $B_R(\frac{1}{2}) = 2$, the uniform error $|\varepsilon(x)| \leq 0.000072$. Adapting the mesh leads to a slight improvement.

2.6.3 Quadrature Methods

Each of the previously considered numerical methods requires the evaluation of definite integrals in order to produce a function that approximates the solution of a Fredholm integral equation on the interval $[a, b]$. Numerical quadrature methods, on the other hand, require the solution of a system of linear equations in order to do so.

The purpose of any numerical quadrature method is to approximate the definite integral of a continuous function $f(x)$ on a closed interval $[a, b]$ with a finite sum. Regardless of the chosen method, the approximation always takes the form

$$\int_a^b f(x)\,\mathrm{d}x = \sum_{i=1}^{n} w_i\, f(x_i) + E.$$

Each quadrature method requires a predetermined set $\{x_1, \ldots, x_n\}$ of *nodes* with $a \leq x_1 < \ldots < x_n \leq b$ and a set of positive *weights* $\{w_1, \ldots, w_n\}$. The error term E depends upon n, a, b, and the value of some higher derivative of $f(x)$ at an interior point of the interval. The choice of method might depend upon the form of $f(x)$, the interval of integration, the amount of computation required to achieve a given level accuracy, or other factors.

Although there are many such quadrature formulas, they fall into two major classes. Newton–Cotes quadrature formulas require equally spaced nodes, so that

$$\Delta x = x_i - x_{i-1} = \frac{b-a}{n}$$

for all $i = 1, \ldots, n$ and a specific set of weights that do not depend upon the nodes. Formulas of this type include the Trapezoid rule, Simpson's rule, and Bode's rule. On the other hand, the nodes in Gaussian quadrature formulas are chosen to be the zeroes of some orthogonal polynomial of degree n, and the weights are given in terms of these polynomials and/or their derivatives evaluated at the nodes. Common

choices are Legendre, Chebychev, or Jacobi polynomials. (Laguerre polynomials are used if the interval is $[0,+\infty)$, and Hermite polynomials are used if the interval is $(-\infty,+\infty)$.)

Now consider the integral equation

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t,$$

where $f(x)$ is continuous on the interval $[a,b]$ and the kernel $K(x,t)$ is continuous on the square $Q(a,b)$. After choosing an appropriate numerical method, we substitute each node x_i into the integral equation to obtain the n equations

$$\phi(x_i) = f(x_i) + \lambda \int_a^b K(x_i,t)\,\phi(t)\,\mathrm{d}t.$$

By substituting the nodes and positive weights indicated for the chosen method, the value of each of these integrals can be expressed in the form

$$\int_a^b K(x_i,t)\,\phi(t)\,\mathrm{d}t = \sum_{j=1}^{n} w_j\,K(x_i,x_j)\,\phi(x_j) + E(x_i).$$

Replacing the definite integrals with the finite sums produces the n equations

$$\phi(x_i) = f(x_i) + \lambda \sum_{j=1}^{n} w_j\,K(x_i,x_j)\,\phi(x_j) + \lambda\,E(x_i).$$

After discarding the error term, we arrive at the system

$$y_i = f(x_i) + \lambda \sum_{j=1}^{n} w_j\,K(x_i,x_j)\,y_j$$

of n equations in the n unknowns y_i. It was necessary to replace the *exact* values $\phi(x_i)$ with the *approximate* values y_i since the error terms were discarded. If the $E(x_i)$ is small, then y_i is close to $\phi(x_i)$. In matrix form, this linear system becomes

$$(\mathbf{I} - \lambda\,\mathbf{KW})\,\mathbf{y} = \mathbf{f}.$$

In this matrix equation, we have set $\mathbf{K} = (K_{ij}) = (K(x_i,x_j))$, $\mathbf{y} = (y_1,\ldots,y_n)^{\mathrm{T}}$, and $\mathbf{f} = (f(x_1),\ldots,f(x_n))^{\mathrm{T}}$. The matrix $\mathbf{W} = (W_{ij})$ is a diagonal matrix with weights w_i appearing on the diagonal, i.e., $W_{ii} = w_i$ and $W_{ij} = 0$, if $i \neq j$. Assuming that the matrix $(\mathbf{I} - \lambda\,\mathbf{KW})$ is invertible, the solution assumes the form

$$\mathbf{y} = (\mathbf{I} - \lambda\,\mathbf{KW})^{-1}\,\mathbf{f}.$$

Once $\mathbf{y} = (y_1,\ldots,y_n)^{\mathrm{T}}$ is determined, a continuous interpolating function $y(x)$ can be constructed on the interval $[a,b]$ that passes through all of the points (x_i,y_i).

A simple way to accomplish this construction would be to define a continuous interpolating function by setting

$$y(x) = f(x) + \lambda \sum_{j=1}^{n} w_j K(x, x_j) y_j.$$

Another simple way to produce a continuous interpolation function would be to use the classical Lagrange interpolation formula

$$y(x) = \sum_{i=1}^{n} y_i \left(\prod_{\substack{1 \le k \le n \\ k \ne i}} \frac{x - x_k}{x_i - x_k} \right)$$

directly when n is small. For example, if $n = 3$, then

$$\begin{aligned} y(x) = {} & y_1 \left(\frac{x - x_2}{x_1 - x_2} \right) \left(\frac{x - x_3}{x_1 - x_3} \right) \\ & + y_2 \left(\frac{x - x_1}{x_2 - x_1} \right) \left(\frac{x - x_3}{x_2 - x_3} \right) \\ & + y_3 \left(\frac{x - x_1}{x_3 - x_1} \right) \left(\frac{x - x_2}{x_3 - x_2} \right). \end{aligned}$$

By inspection, it is clear that $y(x_1) = y_1$, $y(x_2) = y_2$, and $y(x_3) = y_3$.

However, if n is large, then essentially unpredictable anomalies may arise in the interpolation function. For example, if $n = 50$, then it is possible to construct a polynomial $p(x)$ of degree 49 with the Lagrange formula so that $y_i = p(x_i)$ for all $i = 1, \ldots, 50$. Its derivative $p'(x)$ might have as many as 48 simple zeroes, at which as many as 48 relative extrema may occur. If a large number of these zeroes fall between two successive nodes, then it is likely that "polynomial wiggles" would be introduced into the graph of the interpolating polynomial between those two nodes. As a consequence, $p(x)$ would then approximate the solution poorly between successive nodes.

To avoid the introduction of polynomial wiggles, *natural cubic splines* are often used in practice. In a sense, cubic splines are the smoothest possible interpolating functions.

Briefly, a cubic spline is a piecewise cubic polynomial. After a choice of nodes $\{x_0, x_1, \ldots, x_n\}$ is made, with $a = x_0 < x_1 \ldots < x_n = b$, a cubic polynomial $q_i(x)$ is defined on each subinterval $[x_{i-1}, x_i]$, for $i = 1, \ldots, n$. These cubics are defined uniquely by imposing continuity conditions on each $q_i(x)$ and its first two derivatives at the interior nodes and requiring that $q''(a) = q''(b) = 0$ at the endpoints. To be more specific, for each $i = 1, \ldots, n-1$, we require that $q_i(x_i) = q_{i+1}(x_i)$, $q_i'(x_i) = q_{i+1}'(x_i)$ and that $q_i''(x_i) = q_{i+1}''(x_i)$. Since the values of the adjacent cubics and their first two derivatives match at the nodes, the interpolation is quite smooth. Furthermore, since each $q_i(x)$ can have no more than two critical points in each subinterval, polynomial wiggles cannot arise.

Illustrative Examples

- *Example 1*: Consider the inhomogeneous integral equation

$$\phi(x) = x^3 + \frac{1}{2}\int_{-1}^{+1} \sin\left(\frac{\pi xt}{2}\right)\phi(t)\,\mathrm{d}t. \tag{2.30}$$

After replacing the kernel with a separable one derived from fifth partial sum of the Maclaurin series for $\sin x$, we can determine an approximation $y(x)$ to $\phi(x)$ by solving the equation

$$y(x) = x^3 + \frac{1}{2}\int_{-1}^{+1}\left(\frac{1}{2}\pi xt - \frac{1}{48}\pi^3x^3t^3 + \frac{1}{3840}\pi^5x^5t^5\right)y(t)\,\mathrm{d}t.$$

By employing the method of Sect. 1.3, we obtain

$$y(x) = 0.565620x + 0.847692x^3 + 0.014047x^5.$$

This approximation is quite accurate in the sense that the residual

$$\begin{aligned}\eta_y(x) &= y(x) - x^3 - \frac{1}{2}\int_{-1}^{+1}\sin\left(\frac{\pi xt}{2}\right)y(t)\,\mathrm{d}t\\ &\approx 0.000660x^7 - 0.000018x^9.\end{aligned}$$

(a) An application of Simpson's rule yields another approximation $z(x)$ to $\phi(x)$. If we partition the interval $[-1,+1]$ into six subintervals, then the length of each subinterval is $\Delta x = 1/3$. The equally distributed nodes are

$$\{x_1,\ldots,x_7\} = \{-1,-\tfrac{2}{3},-\tfrac{1}{3},0,+\tfrac{1}{3},+\tfrac{2}{3},+1\},$$

and the corresponding weights are

$$\{w_1,\ldots,w_7\} = \{\tfrac{1}{9},\tfrac{4}{9},\tfrac{2}{9},\tfrac{4}{9},\tfrac{2}{9},\tfrac{4}{9},\tfrac{1}{9}\}.$$

If we set $x = x_i$ in Eq. (2.30), then we obtain

$$\phi(x_i) = x_i^3 + \frac{1}{2}\int_{-1}^{+1}\sin\left(\frac{\pi x_i t}{2}\right)\phi(t)\,\mathrm{d}t$$

for each $i = 1,\ldots,7$. Approximating each integral with Simpson's Rule yields the approximate equations

$$z(x_i) = x_i^3 + \frac{1}{2}\sum_{j=1}^{7} w_j \sin\left(\frac{\pi x_i x_j}{2}\right)z(x_j).$$

Now set $\mathbf{f} = (-1, -\frac{8}{27}, -\frac{1}{27}, 0, +\frac{1}{27}, +\frac{8}{27}, +1)^{\mathrm{T}}$, and solve the linear system

$$(\mathbf{I} - \tfrac{1}{2}\mathbf{KW})\mathbf{z} = \mathbf{f},$$

where $\mathbf{K} = (K_{ij}) = (\sin(\pi x_i x_j/2))$ and $\mathbf{W} = (W_{ij}) = w_i\,\delta_{ij}$, to obtain

$$\mathbf{z} = \begin{pmatrix} -1.425492 \\ -0.630291 \\ -0.220551 \\ 0.000000 \\ +0.220551 \\ +0.630291 \\ +1.425492 \end{pmatrix}.$$

After substituting these values into the interpolation formula, we obtain the approximate solution

$$\begin{aligned} z(x) &= x^3 + \frac{1}{2}\sum_{j=1}^{7} w_j \sin\left(\frac{\pi x x_j}{2}\right) z(x_j) \\ &\approx 0.567809\,x + 0.842899\,x^3 + 0.015578\,x^5. \end{aligned}$$

In this case, the residual is

$$\begin{aligned} \eta_z(x) &= z(x) - x^3 - \frac{1}{2}\int_{-1}^{+1} \sin\left(\frac{\pi x t}{2}\right) z(t)\,\mathrm{d}t \\ &\approx 0.0023436\,x - 0.004890\,x^3 + 0.001542\,x^5 \\ &\quad - 0.000158\,x^7 + 0.000007\,x^9. \end{aligned}$$

(b) A comparable approximation to the solution $\phi(x)$ can be had by employing Gaussian quadrature.

The Legendre polynomial of degree seven is

$$P_7(x) = \frac{1}{16}\left(429x^7 - 693x^5 + 315x^3 - 35x\right).$$

With this method, the nodes $\{u_1, \ldots, u_7\}$ are the seven roots of $P_7(x)$

$$\{-0.949107, -0.741531, -0.405845, 0, +0.405845, +0.741531, +0.949107\},$$

and the corresponding weights $\{v_1, \ldots, v_7\}$ are computed from the formula

$$v_i = \frac{2}{(1-u_i^2)\,(P_7'(u_i))^2}$$

to be

$$\{0.129484, 0.279705, 0.381830, 0.417959, 0.381830, 0.279705, 0.129484\}.$$

If we set $x = u_i$ in Eq. (2.30), then we obtain

$$\phi(u_i) = u_i^3 + \frac{1}{2}\int_{-1}^{+1} \sin\left(\frac{\pi u_i t}{2}\right)\phi(t)\,\mathrm{d}t$$

for each $i = 1, \ldots, 7$. Approximating each of these integrals by Gaussian quadrature yields the approximate equations

$$g(u_i) = u_i^3 + \frac{1}{2}\sum_{j=1}^{7} v_j \sin\left(\frac{\pi u_i u_j}{2}\right) g(u_j).$$

Now set $\mathbf{h} = (u_1^3, \ldots, u_7^3)^{\mathrm{T}}$ and solve the linear system

$$(\mathbf{I} - \tfrac{1}{2}\mathbf{KV})\mathbf{s} = \mathbf{h},$$

where $\mathbf{K} = (K_{ij}) = \sin(\pi u_i u_j/2)$ and $\mathbf{V} = (V_{ij}) = v_i\,\delta_{ij}$, to obtain

$$\mathbf{s} = \begin{pmatrix} -1.271807 \\ -0.768011 \\ -0.286296 \\ 0.000000 \\ +0.286296 \\ +0.768011 \\ +1.271807 \end{pmatrix}.$$

After substituting these values into the interpolation formula, we obtain

$$\begin{aligned} g(x) &= x^3 + \frac{1}{2}\sum_{j=1}^{7} v_j \sin\left(\frac{\pi x u_j}{2}\right) g(u_j) \\ &= 0.565420\,x + 0.847750\,x^3 + 0.014041\,x^5. \end{aligned}$$

The approximation $g(x)$ is very close to $\phi(x)$, since the residual

$$\eta_g(x) = g(x) - x^3 - \frac{1}{2}\int_{-1}^{+1} \sin\left(\frac{\pi xt}{2}\right) g(t)$$
$$\approx 10^{-8} \times (-4.337x + 1.543x^3 - 0.181x^5).$$

The approximation $g(x)$ is also very close to $y(x)$ determined above as the solution to the integral equation with the separable kernel.

The graphs of the approximations $y(x)$, $z(x)$, and $g(x)$ are "within the ink" of each other, i.e., if all three approximations are plotted in the same plane, they are indistinguishable.

The number of nodes in the illustrative examples was deliberately chosen to be small in order to avoid unnecessary numerical complications. It should be noted that significant problems relating to stability may arise if the number of nodes and/or the number of basis functions is increased.

Springer

http://www.springer.com/978-0-8176-8348-1

The Classical Theory of Integral Equations
A Concise Treatment
Zemyan, S.M.
2012, XIII, 344 p., Hardcover
ISBN: 978-0-8176-8348-1
A product of Birkhäuser Basel